MODERATING PERFORMANCE ASSESSMENT: COMPUTER OR CONNOISSEUR? COMPARING OBJECTIVE AND SUBJECTIVE MEASUREMENTS OF VIOLIN PROFICIENCY

Patrick Henry Early

A thesis submitted in partial fulfilment of the
requirements of the University of Lincoln for the
degree of Doctor of Philosophy (Professional)

January 2018

Abstract

Objectivity in music assessment draws attention to milestones in learning, and creates a framework for the determination and implementation of set levels of attainment. In so doing, it creates a space where assumptions, which obscure long-term pedagogical objectives, can be challenged. Through innovations in technology, an array of 10 variables was used to observe novice violinists for the presence of verifiable milestones in string learning. The results were scored quantitatively and compared to subjective grade results, to test a hypothesis which postulates a link between them. Variables, under headings such as pitch discrimination, intonation and rhythm accuracy, posture angles, tone production and sight reading abilities, create the profile of the player from which a data rich analysis is made. The accumulated scores from these variables are compared to conventional subjective outcomes of grade examinations. The analysis found correlations between objective and subjective methods, bringing to the surface an innovative approach to formative performance music assessment.

Statement of Authenticity

I certify that this thesis is my own original work. All data gathered in the study has been used for this research alone. I have acknowledged clearly any materials from other sources adopted in the study.

Patrick H. Early

Acknowledgements

I would like to acknowledge the help and support I have received from all the faculty at Lincoln University. In particular, to Prof Terence Karran, whose guidance, expertise, and good humour, has made the journey so much more enjoyable. His unwavering support, belief in my work and sustained commitment to helping me realise my goals, particularly at times when I might otherwise have given up, must be acknowledged here. Thank you Terence, I am truly grateful. Thanks also to Dr Joss Winn for preparing and facilitating the study school programs both in Lincoln and Mayo which helped to bridge the many learning gaps along the way. Thanks also to Nettie for minding the animals while I was away.

Dedication

To the memory of my brother Jimmie who was a wonderful musician and is sadly missed. To my Dad, who gave me my first violin, and still takes a keen interest in music. To my two girls, Hannah and Alannah, and to my four grandchildren, Jack, Maddie, Ryan and Conor who are just starting out on their journey of learning and discovery.

Publications

Early, P., (2018) Introducing a culture of string playing in primary school: A teachers guide. *South East Regional Research Journal VI.* Kildare: Outside The Box Learning Resources Ltd.

Early, P., (2017) Music education, a transformation at primary level. *South East Regional Research Journal V.* Kildare: Outside The Box Learning Resources Ltd.

Early, P., (2016) *Primary School Strings, Books 1-11: teaching material for junior infants – 6th class string learning.* Westport: Jules.

Early, P., (2010) *Reading for D Whistle*: teaching music literacy in primary school. Unpublished manuscript.

Early, P., (2008) *Metronome Tuning*©: temperament setting using interference waves of harmonics. Unpublished manuscript.

Early, P., (2003) *The development of the violin treatise over three centuries*. M.A. dissertation. University of Limerick. Ireland.

Early, P., (2000) *The Fingerboard Sliderule* ©: a method for teaching finger patterns on the violin fingerboard. Unpublished manuscript.

Early, P., (1999) *Absolute Beginners*: introduction to 4-part string performance. Unpublished manuscript.

Contents

List of Figures

1 Chapter One Introduction

Machine enabled assessment for novice violinists is becoming increasingly possible with information and communication technologies, which are helping to focus attention on aspects of learning that are finite and measurable, similar to developments already taking place in the sports sciences. In this thesis, a novel assessment tool is put forward which exploits observable traits relating to comprehension, posture and sound analysis during the string learning process, generating machine measurements which are verifiable and repeatable. These developments enable accurate determinations to be made in the field environment as the technologies are portable, immediate, flexible and adaptable. Combining these technologies has helped to create a unified objective assessment model. The study explores this approach in detail, and compares objective outcomes to conventional methods of novice string performance assessment, which rely on a mostly subjective interpretation, often referred to as a non-instrument-specific 'musical outcome'. The study draws on Educational Bildung theory to support pedagogically relevant string learning practices and material. The correlation between objective and subjective methods is demonstrated.

1.1 Thesis Overview

The thesis is broken down into five main sections. The introduction section sets out the goals, objectives and overall aspirations of the research. It lists research questions, which have prompted the research and the hypothesis statement. This is followed by the literature review section, which looks closely at other research in this area, including aspects of performance assessment evolution, pitch determination approaches, temperament, intonation and tone production, along with sections on posture and reading, and educational theorists. Following this, the methodology chapter goes into detail about the research design and the processes followed in collecting the data. The data analysis chapter describes aspects of discovery and revelation, concluding with a summary, recommendations, and suggestions for future research.

The research is structured in nature and quantitative in its design. Factual statements, generated because of substantive explanatory theory, are presented to support a hypothesis statement which postulates a link between observational findings of children's string learning and attainment, and grade results (Punch, 2006). Theory generation at a more abstract level, resulting from the processes and technologies employed, constitutes a contribution to knowledge in string learning. This theory verification study explores string performance measurement by incorporating new technologies which shed light on abstract constructs relevant to learning, which are otherwise difficult to define, verify and replicate.

The research undertakes five separate observation sessions with 37 children. The observations measure predefined attainment goals in musicality. These measurements are described and examined for correlations with grade results to support or refute the hypothesis which postulates pre-existing links between them. Ten measurement instruments have been developed to provide data scores for ten variables. These include:

1. Frequency change detection
2. Intonation
3. Rhythm accuracy
4. Angle of instrument
5. Orthogonal bow and elbow angles
6. Bow distance travelled
7. Bow guidance
8-10. Sight-reading accuracy
These are compared individually and collectively with variable 11, which is the Grade result.

Parametric markers of individual string learning produce useful short-term formative assessment feedback, whilst meeting, at the same time, long term pedagogical objectives relating to ergonomics and musical literacy. Explicit goals in the short term, are generally misunderstood for their usefulness because process (the learning) tends to be overshadowed by product (the performance). Shifting focus away from product in assessment, to processes (which make the product possible) is an important consideration in this research. Supported by best pedagogical practice, and reinforced by the latest developments in technology, the research presents a more contemporary and comprehensive approach to assessment and feedback.

1.2 Why this study?

Calls for improved standards of assessment in core curriculum music (Fischer, 2009) are indicative of the need for greater accountability in class performance music education. Popular assessment strategies have been found to be 'non-musical' in nature (McQuarrie and Sherwin, 2013). Lehmann (2014) found the relationship between entrance and final examinations results in music to be moderate. Elliott (1995) suggests that the way curriculum development takes place is inappropriate for music. Butke (2014, 23-27) refers to a "vehicle where both concrete musical behaviours and expressivities can be assessed", and where "emphasis is on process and the kinaesthetic analysis of musical concepts as opposed to being a performance product" satisfies the need for outcomes to be observable and measurable, using an assessment rubric. Mark (1982) indicates that the abundant justification for the existence of music education warrants a more robust assessment model.

Music "forms the backbone of cultural participation in adulthood" (Abeßer et al., 2013, 1). Cultural participation in adulthood has profound implications for how society functions and accommodates the social development of its citizens. Many parallels between subject disciplines and outcomes in later life draw similar claims. For example, Mathematics informs a technologically literate community. Literature is key in communicating ideas. The anomaly arises, however, when one considers the contrasting randomness with which music is taught and assessed in comparison to other subjects, despite the likelihood of the claim by Abeßer et al. above being factually correct. The role of the evaluator, in making determinations about music performance and the subsequent formative feedback it embodies, is brought into question, as the unevenness of experience among evaluators makes determinations less significant (Byo and Brooks, 1994).

Ambiguities surrounding music education are diverse. An attempt is made in the thesis to introduce clarity, by way of presenting task objectives in a developmental approach to group string teaching, learning and assessment. Time-honoured methods, normally associated with one-to-one teaching, have been adapted and extended to consider group dynamics, reworking existing methods to take account of this.

Comparing scores from objective observations, with conventional subjective grade results, has supported the existence of a link between them. Ultimately, the objective of the research is to streamline the learning experience collectively. Stakeholders in music education can benefit from an observational process which is reflexive, innovative, technologically robust, immediately informative, developmentally based and post-positivist in nature. Administrators, specialists, generalists and students alike can act directly on the observation feedback data in preparation for performance and grade exams. The aspirations of this research are in line with Nielsen's view in relation to *Bildung*.

> "The tasks are (1) To understand better, more thoroughly, and with more certainty what happens in music teaching and musical learning processes; (2) To be able to say how future music teaching and musical learning processes can take place; (3) To contribute in such a way that music teaching and musical learning processes can be realized in a more responsible, well-founded, and well-considered way. A key aspect of this is that music pedagogy as a science must necessarily supersede the current situation of music teaching and learning." (Nielsen, 2009, 32)

1.3 My personal interest in the study

I am a professional musician with a Master of Arts Degree in Classical String Performance. I have come to music education from a performance perspective, with the intention of maintaining and improving standards for children coming to string learning. I have designed many innovative teaching materials, and written several books for music instruction, including a series being used by many of the participants of this study, prior to taking their grade examinations. I work in a private capacity, teaching one-to-one, and in a single whole school music literacy/performance programme. The opportunity to conduct the research was made possible, therefore, by my familiarity with children learning to play the violin.

I relate, as a practitioner and observer over the past 3 decades, to string learning. My interest in the topic, in this instance, is concerned with how standards and teaching strategies, normally available in one to one scenarios, can be applied to the group or classroom environment, and how formative feedback can be generated and returned to the learner and teacher. The research is framed in the context of encouraging students, both private and group, to manage their own progress and take responsibility for what might be described as 'insider' or specialist knowledge. A brief background to the Irish context will help to frame my personal interest in the subject and how it has influenced and guided the study.

The potato famine of 1845 described in Fleischmann (1998), has contributed to an essentialist view of Irish cultural identity, post the founding of the state, and served to embolden a scepticism about 'other' music. Other nations, meanwhile, have embraced the role of the European art music tradition in building a concept of society (Finland is a good comparative example). The delivery of a conventional music education program which includes literacy and expert tuition is uneven, therefore, particularly in the school system, with varying levels of expertise, commitment and expectation being present, from region to region.

This was brought into focus by a survey of school music programs in Ireland commissioned by the Arts Council in 1985 which found that: "The young Irish person has the worst of all European musical worlds" (Herron, 1985, 41), giving rise to a series of reports and studies seeking different explanations for this. The Music Education National Debate (MEND) report by Heneghan (2002) stands out amongst these studies for its examination of the complexity of the problem, exposing the dichotomy between European Art and traditional Irish approaches to music education. Trying to strike a balance between these two views, that is, preserving a traditional ethos and delivering access to a non-traditional repertoire, has been a professional and personal challenge for me as a music teacher.

Complicating further the above dichotomy, Gould (2009, 875) suggests that music is "a commodity sold to pre-service and in-service music teachers. Like all mass-produced consumables, it is valuable to the extent that it is not creative, that is, to the extent that it is reproducible". She is questioning the value of formal art music instruction, as it is primarily concerned with *recreating* works of art, rather than creating them. In addition, Said (2008) argues that piano players who operate at the highest level tend to play a limited classical repertoire and, in so doing, are also being uncreative. These views challenge from a different place, advocacy for the defence, cultivation and promotion of methods of teaching music which have been passed from generation to generation, and from teacher to student in studio and conservatory settings.

It can, therefore, be said that divergent and valid orientations to music education must be challenged in a broader sense, to appreciate fully the lack of consensus in relation to what will be accepted or embraced in the music classroom. Scepticism about pedagogically valid processes, whether rooted in historical resistance to them, or emerging from societal trends towards modernity through advertising and multi-media influences, creates ambivalence which undermines children's abilities to participate fully in a music education program. In my work, this unevenness has prompted the adaptation of a scientific method of determination, as a way of navigating discursive oppositions. This has been achieved by looking chiefly at instrument specific singularities relating to sound, posture, intonation and literacy, supported by pedagogically valid material.

These aspects are a distillation of important elements familiar to insider or music specialists which contribute to a musical outcome at a later stage. Returning feedback about these learning milestones to where they are most needed – the site of learning in the classroom or conservatory, is at the heart of the study. This aspiration resonates with Educational Bildung theory and frames the study.

1.4 Research questions

There are three research questions being addressed in the research study concerning correlation, educational Bildung, and predictability. The first is:

- Does an instrument specific approach to performance assessment correlate with musical outcomes in practical ABRSM grade examinations?

The section dealing with the Associated Board of the Royal Schools of Music (hereafter described as ABRSM) rubrics illustrates how the grade system, by design, makes determinations about ability, based on what it describes in its literature as 'musical outcomes'. This largely subjective appraisal would confirm that, as is indicated in the examiner role and person descriptors discussed on page 50, the examiner, while being a music expert, is not required to be an instrumentalist in the discipline being examined. So this question asks if attributes, deemed to be essential by a music specialist in the instrument discipline being examined, can be measured, scored and compared to conventional grade outcome results. This question is at the heart of the research and central to the observation study. The second question is:

- How can educational Bildung inform string learning?

The instructional material deemed essential for string learning, and the established canon of works, are being replaced with material which represents wider popular and anthropological orientations to music making. According to Bildung theory, however, the ability of music being learnt to affect the potential for further learning to take place is under threat, as the lines between emotion and cognition become blurred. The research looks at this point to frame the approach and justify both literacy, and a music learning process which is grounded in sound pedagogical theory. The final question is:

- To what extent can an objective observation process predict a subjective 'musical outcome' product, and what are the latent underlying predictors behind the ten observational variables?

The analysis section looks at correlation, regression and causal path analysis to better understand the statistical findings of the observation study comparison, between objective and subjective methods of assessment.

1.5 The hypothesis

The null hypothesis states: Relationships between objective observation scores and subjective grade results are random. Ho $p=0.5$. The alternative hypothesis asserts that there is a link between the two. H1 $p\neq0.5$

Through a series of objective observations, and the scores they generate, it is postulated that predictions can be made about attainment in conventional music performance examinations. The logic of the theoretical construct, underlying the hypothesis postulates that the presence of skills being explored in the ten parametric observation tests which make up the study, signal important underlying abilities. These abilities are foundational requirements on which the means to succeed in grade examinations can be tested. Conversely, it is postulated that the absence of these indicators, evidenced through the variable scores, makes attainment less likely.

Correlations between levels obtained in these constructs and subjective grade outcome score levels are examined. The validity of claim is tested to a *P*-value level of <0.05. The alpha level of significance on which the *P*-values threshold is measured against is also 0.05 reducing the likelihood of a type 1 error to a one in twenty chance. A type 1 error occurs “when we believe that there is a genuine effect in our population when in fact there isn’t” (Field, 2002, 784,). The testing process is as follows.

1/ The test for **normal distributions** of means in each variable.

Testing for normal distribution serves to eliminate the possibility of outliers distorting the sample distribution profile. For instance, students with hearing difficulties would create a platykurtic component in the probability distribution of a pitch discrimination test, diluting the power of the sample to represent the population of students with normal hearing being observed. The standard deviation of the sample data is calculated in the following manner.

$$s = \frac{\sqrt{\sum (x - \overline{x})^2}}{\mathrm{n} - 1}$$

2/ The tests for **the degree of association between variable x and variable y** using product-moment correlation coefficients are calculated thus:

$$r = \frac{\sum xy - n\overline{x} \text{ x } \overline{y}}{\sqrt{(\sum x^2 - n\overline{x}^2)(\sum y^2 - n\overline{y}^2)}}$$

3/ Within this thesis, the **level of significance** is set at 0.05. The confidence level indicates the certainty of the margin of error. It is expressed as a percentage and represents how often the true percentage of the population would pick an answer that lies within the margin of error. With a 95% confidence level, there is 1 chance in 20 that we would get a false positive result. In this analysis, the significance level for the tests was set at 5%; however, where the significance of a statistical test is greater than 5% (for example, where it is 1%), this will be reported. The decision to use a 5% (as opposed to a 1% or 10% significance level) is arbitrary but, as Gall et al. (2007) and Cowles and Davis (1982), report, a 5% significance level is invariably used in studies of this kind, and across the social sciences.

4/ The **t test for difference** from 0 (where the null hypothesis states that the relationship is 0) is calculated by

$$t = \frac{\sqrt{\sum n - 2}}{\sqrt{1 - r}}$$

The **null hypothesis** of no correlation should be **rejected if** *t* is large, positive or negative (linear –1 or +1)

1.6 What the study hopes to achieve

Historical musicology is being supplanted by a range of interdisciplinary approaches which dilute the focus on masterworks which academies teach musicians to perform. While the inclusion of 'other' music may be an important feature of repertoire preparation, with ethnic and popular music competing for inclusion in the school system, material without a pedagogical basis creates a gap in instructional continuity. The research hopes to fill this gap by making explicit a cause-and-effect relationship between learning objectives and grade outcomes. We are reminded that "studying music is no less 'scientific' than studying science itself" (Dunsby, 1995, 14).

The study defines an array of constructs which meaningfully measure common learning objectives in string performance instruction, in which artistry is rooted. It is hoped these measurements, and their analysis, will offer valuable feedback to students throughout the learning process. Sloboda (1985, 234) advises that assessment "should be taken in conjunction with other evidence," given that emotion and motivation described by McPherson and O'Neill (2010) are out of step with precision. However, the knowledge created, concerning how observation parameters correlate with each other, and respective grades, will be of interest to music educators, curriculum designers and music students.

It is hoped that a structure is created to help inform pre-service generalist teachers, that is, teachers studying to become teachers, who have not yet qualified. The study will make explicit for them ways of understanding stages and milestones in string learning, particularly regarding the long-term view, and the processes involved. Informal assessment should be happening continuously in the music class, according to Millican (2015), therefore an ability to understand and log important learning milestones must be of concern in an ongoing basis. Documenting what specifically improves is important, according to Pellegrino et al. (2015), and to do this there must be ways of observing particular singularities and knowing how to assess them. Making explicit the stages of learning, and how they relate to outcomes, provides a rationale through which generalists can better understand essentially vague processes, rather than being constrained by the need to produce a performance product which demonstrates school music output, as is often the case.

Two new technologies relating to machine measurement have been identified and incorporated into the study (and will be described in detail later in the thesis). Melodine software was used to measure intonation accuracy, and Kinovea was introduced to assist with determinations and measurements relating to posture, bow movement and bow angles. These innovative approaches to performance assessment, coupled with a literacy aptitude component, form the basis of the Primary Observation Package (hereafter designated as POP) developed uniquely for this research study. In addition to these, other software used in the study includes Sibelius, which was used for typesetting the sight-reading test, the rhythm test and for typing and playing the pitch perception variable questions. GarageBand recording software was used for recording student responses to the rhythm accuracy test, also discussed in detail in the methodology chapter.

1.7 Other studies

There are arguments for, and against, different modes of assessment in music education. Fine arts teachers in the United States favour non-achievement criteria, such as attendance and participation, over subjective methods to rate students. Russell and Austin, (2010) advocate assessment tools as a more valid and ethical way of determining student achievement in string playing. Pellegrino et al., (2015) suggests rating scales, which list specific learning criteria in string playing, are useful, arguing that rating scales direct student focus, when developed into summative reports. Asmus (1999) states that future learning experiences can be identified by measuring acquired knowledge objectively. Cangro (2016) recommends an interactive, student-centred, standards-based learning environment. Kohn (2000) argues that grading adversely affects motivation, suggesting that an agenda of ranking, rather than rating, may be present, and cites a socioeconomic explanation for results. He suggests that a more accurate way of making assessments would be to look at the level of cultural capital imbedded within the learning dynamic, through the affluence and social status of the learner. Richmond (2002) indicates how the assessment process has been complicated, with incidence of disputes over grading. Such disputes shift attention further away from measuring standards.

Non-musical tests by Rickard et al. (2012) give indications of overall abilities and predispositions to learning. The Children's Memory Scale (CMS) by Cohen (1997), the Kaufman (2004) Brief Intelligence Test (KBIT-2), the Culture-Free Self-Esteem Inventories (CFSEI-3) by Battle (2002), the Rosenberg (1965) Self-Esteem Scale (SES) and the Social Skills Rating System (SSRS) by Gresham and Elliott (1990), all offer valuable insights into student potential.

The lack of standardised observation instruments to measure musical abilities, and the lack of interest in psychometrically sound tests capable of measuring differences described by Law and Zentner (2012), inspired an assessment design which tries to uncover hidden realities about student potential. Law was prompted by concern with how untrained musicians with potential, and mediocre ones (without potential), go undetected following considerable training, because of assessment design shortcomings. In relation to this, the Profile of Music Perception Skills (PROMS) test was cited as providing "researchers with an instrument to assess the level of listeners' perceptual musicality objectively" (Law and Zentner, 2012, 11). Referring to subjects as potential 'musical sleepers' or 'sleeping musicians', and articulating a difference in causality, the PROMS test attempts to anticipate musicality potential before training begins, stating that talented individuals can miss opportunities for development, while others who receive a lot of training are given tools to mask inherent inabilities in their playing. Gordon’s Advanced Measures of Music Audiation (1990) was used for convergent validity to establish internal consistency with the PROMS test.

Law and Zentner (2012, 2) refer to objective assessment having a “role in devising treatment plans” and identifying individuals whose ability is more or less than expected, based on training given, in order to reduce categorisation errors. Proposing a culturally divergent approach, and noting an absence of diatonic key relationships in modern Western, African and Indian music cultures, the term ‘culturally evolved musical system’ articulated this aim in their research design.

The understanding of musical ability in that instance was fashioned in the context of ‘potential’ prior to training according to Shuter-Dyson & Gabriel (1981). Law and Zentner, (2012) also suggests that there is no agreement about how musical ability might be measured with objective tasks, arguing that validity and reliability in the past have been tenuous in comparison to contemporary standards. This approach to assessing musical objectives objectively, while being an update on previous studies, did not address what it referred to as the measurement of music production abilities (i.e., playing the instrument). This gap in objective assessment is the focus of this study, in that it sets out to map instrument specific abilities, both with regard to perception and aptitude, in novice players.

While analysis of sound alone, as described by Charles (2010), bow tracking methods described by Pardue et al. (2015) and error detection methods detailed in Luo et al. (2015), offer a way forward, objective approaches such as Wu et al. (2016) coupled with an understanding of the incremental nature of string learning discussed in Pellegrino et al. (2015) resonate particularly with the research design and approach taken. This is because cognitive overload can obscure the potential of analysis to separate out complex layers of learning, and isolate the ones which need attention.

1.8 Subjective assessment

The first syllabus of the U.K.'s Associated Board of the Royal Schools of Music

(ABRSM) appeared in 1890, with 1,141 candidates taking performance music examinations. Today, more than 650,000 take ABRSM examinations annually. The ABRSM provides stimulus for high standards in music performance, with grade examinations based on the performance of three pieces, along with selected scales and arpeggios, sight-reading, and aural tests. The ABRSM graded music exams provide a structured framework for progression from beginner to advanced musician. Requirements and performance criteria are established in their syllabus and renewed every three years , see Scaife (2015). The assessment and determination of grades by the ABRSM is largely subjective, and is based on the notion of learnt expertise and connoisseurship. Underlying the latter is the assumption that "growth in taste and appreciation has been held to be correlative with growth in musical skills, knowledge and the ability to comprehend and discriminate the musical qualities" (Broudy, 2008, 202).

The arrival of new technologies which can gather and analyse highly accurate audio data and motion capture, has enabled a reappraisal of the process of musical assessment in the 21st century. The arrival of such technologies not only paves the way for a more scientific verification model of assessment of musical abilities, but also signals a paradigm shift in the way students learn music.

This study compares a scientific objective method of assessment, which utilises these technologies, with the more conventional subjective appraisal, which relies on consistency and evenness in connoisseurship, as utilised by the ABRSM. For this research, the primary observation package (POP) was designed to gather numerical data, creating an empirical profile of student learning milestones and attainment, deemed relevant to string playing. This scientific measurement package, which removes the vagaries of aesthetic considerations by default, consists of 10 variable constructs and measures five domains of learning, supported by string learning theory.

The observational research consists of the scientific measurement of data gathered by the observation of students learning to play the violin. For example, it is possible to tell, to a high level of measurable accuracy, whether a student is playing a note in tune, or sharp, or flat, and by how much. This type of numeric data forms the basis for the variables used to operationalise the study. A definitive objective assessment, indicating level of student attainment, was constructed from an analysis of the data, completing the first stage of the research. Secondly, a comparison was made between this objective assessment, and results which were obtained from conventional ABRSM examinations, taken shortly afterwards.

1.9 About the school

The Irish school where the research took place is an all Irish speaking school and had 206 students (boys and girls) attending. Music was an integral part of school activities with children being encouraged to participate in playing whistle, accordion, and fiddle instruments in traditional music sessions and in a half hour group lesson, provided for each class each week, for specialist instruction string playing and music literacy. The whole school program was supported by parents and teachers, and it was decided to set an attainment level of Grade II as an achievable target goal, following 8 years of group string tuition.

While all students participated in the group class at junior infant, senior infant, 1^{st}, 2^{nd}, 3^{rd}, and 4^{th} class level, by the time students were in 5^{th} and 6^{th} class, many had dropped out, with an average of only 50% remaining by graduation and less than 25% of these taking grade examinations. This was deemed to be acceptable by the school, as there were many alternative options provided for the students to take up, including traditional musical activities. Half of the participants in the study came from students in the 5^{th} and 6^{th} classes, taking Grade I and Grade II examinations respectively, and the remainder were made up of private students who studied outside school hours. The observations, from which the variable determinations were made, took place at the school. The process and observation schedule is covered in the section on methodology on page 75.

Figure 1. Group String Rehearsal

The above group panorama photo was taken during a rehearsal for the official opening of the school which was attended by the head of government at the time, Prime Minister Enda Kenny. As can be seen, given that each child in the photo has a violin, the problem of making determinations about student ability is a formidable challenge, considering time constraints and the numbers involved. This is compounded by a scepticism about music examination as a process, with many students and parents preferring music participation to be exclusively for their own 'enjoyment'.

Attempts to improve standards, therefore, gave rise to an informal objective assessment model. That is, a way of determining if improvements were taking place without students feeling that they were being examined. This was considered on the basis that many determinations can be made objectively, with the help of technology. The observation criteria follow established string teaching goals and measurable learning milestones, but are conducted in an informal way. Furthermore, the Primary School Curriculum (NCCA, 1999) recommends that *all* subjects be assessed. Indeed, the Irish National Council for Curriculum and Assessment states: "Observation helps the teacher to find out the varying degrees of success with which a child acquires and masters different skills and knowledge and then to adjust teaching and learning contexts accordingly" (NCCA, 2007, 46). The study presented here demonstrates a process through which this aspiration can be realised in music education, almost 20 years after these recommendations were first published.

1.10 Motivation for the study

Improving methods of string assessment has been the main motivational factor behind this study. In 2011, a whole school music education programme for strings was undertaken by the school where the research took place. Addressing developmental learning needs within a cultural context, but without compromising long-term string learning objectives, was paramount in the design of the programme. Students who chose to do so, could reach the equivalent of a Grade II ABRSM in performance standard at the end of the primary school cycle, giving them the possibility of maintaining continuum at second level (Grade V at Junior Certificate – Grade VII at Senior Certificate), making comparative performance criteria in literacy and string playing technique at third level entry a real possibility.

Each class, from junior infants upwards, was given specific learning objectives to follow within a tasks list, drawn up for both music students and generalists. Children are prepared for the specialist music class in advance, during the normal school day. A series of books, designed and published for each of the eight primary school years, was provided, with whiteboard projections to facilitate continuity and parity with normal classwork. The books were in landscape format, and small enough to fit into the child's small violin case (A4 black and white portrait music pages are impractical for teaching small children). Each child was provided with a newly-sized violin each September.

The ABRSM examine student music performance in 93 countries worldwide and issue a new syllabus of requirements and pieces every three years (the ABRSM is considered in more depth in section 2.7, below). The material encompasses medieval, baroque, classical, traditional and contemporary music genres. Scales, sight-reading and aural tests also feature in the grade examinations. This approach to string learning uses a subjective method of assessment, with the examiner making determinations on the basis of a practical examination. This provided a useful framework from which a comparison with objective methods using the new technology could be made.

Prior to students taking examinations, observational sessions with 5th and 6th class students, and private students were undertaken. Constructs mentioned earlier, objectively measured string performance under ten separate headings. These scores were compared later to grade outcome results, and found to have significant levels of correlation. At this point, the predictive value of the observations was realised, and the potential for using them during the learning process became apparent.

Putting a school music programme together, despite the many difficulties, is a relatively straight-forward logistical undertaking. Children are incentivised by working together, unlike the one-to-one model. The challenge begins, however, when one must evaluate or validate the learning which has taken place. Here, the emerging technology has been exploited to open up a new conversation between teacher and student, about gaps which exist in the learning. The class teacher plays the crucial role in determining the success of a school music programme, and is assisted by long-term milestone predictors, newly evidenced by the technology.

Testing the hypothesis, which postulated a link between objective and subjective methods of assessment in novice string performance, supported the link. A correlation was established between both approaches, at the 0.05 significance level or higher. In other words, potential, as defined by the variables, was found to translate directly into musical ability in performance. The study introduced a learning goals structure for generalist classroom teachers and a primary observation package for measuring student attainment. Given the positive correlation evidenced in the data in Chapter 4, high levels of reliability and verifiability have been demonstrated supporting instrument specific approaches, which, furthermore, deliver formative feedback. In doing this, long-term, process-orientated, pedagogical objectives can be shown to advance short-term performance concerns. This demonstrates how instructional continuity gaps discussed throughout, relating to posture, tone production and literacy, can address standards achievable in the 'musical outcome' based grade examinations.

Threshold concepts describe a phenomenon which takes place when the significance of key concepts become apparent, see Meyer and Land (2003). Ideally, all learning should entail elements of such a transformation, not only on the part of the learner, but also on the part of the evolving teacher. Learning is not a circular movement of fixed or 'banked' knowledge, but rather an exponentially developing system of exploration and transformation. Measurement of milestones in string learning, and the ability to understand and assimilate what was previously 'insider' knowledge, constitutes the crossing of such a threshold.

The emerging use of technology, by Ng et al. (2007) for example, sheds light on this type of abstract 'insider' knowledge previously unavailable to the generalist. Objectivity in performance music assessment in Wu et al. (2016) can be adapted to draw attention to learning milestones which take place in classroom scenarios. Observational determinations which measure progress of children's technique and related markers of ability to play and read music, can be returned directly back into the learning domain. Despite running the risks of what Smith (2011) refers to as 'metricophilia', or an over-reliance on statistics, qualitative assumptions in music education run a similar risk of being distracted by identity formation issues, which have more to do with anthropology than pedagogy. A better balance is struck, as the technical means to measure tenuous or abstract events, such as ones described by Dittmar et al. (2012) become more reliable.

The variables used in the study (for which data was derived using information and communication technologies), are pitch discrimination, intonation accuracy, rhythmic precision, posture angles, tone production and sight-reading abilities, and are firmly grounded in music education theory. Measurement of learning in these areas creates a profile which makes data-rich analysis possible. The hypothesis postulating a link between the variables has been supported. As a result, this supply of new knowledge empowers the pre-service generalist to respond much more effectively. Measurement of learning in these areas has been the main motivation for the study.

1.11 Application of the study thereafter

Nielsen (2009) describes how the relationship between music research and its application should be both remote and close up at the same time, in order for it to be relevant and objective. Potentialities made accessible through the verbalisation of thought is in keeping with Humboldt's vision for pedagogical research to be applied directly back into the learning environment according to Morgan (2011). The term *Bildung*, discussed in section 2.8, embodies this central concept, whereby learning is used to enhance future learning. The research study at hand can be usefully applied directly back into present-day learning environments.

Take, for instance, a young primary school teacher who is struggling to deliver a music instruction programme effectively, to children in his/her care. He/she may have limited ability in performance on an instrument, and may be feeling vulnerable and compromised in the capacity to offer a basic pedagogical system of instrument instruction to students, while keeping pace with set curriculum constraints. This is compounded by the fact that a 'letters' system (where instructional pages consisting entirely of letters only, representing note pitches for given tunes, are provided instead of music notation) to teach tunes 'aurally' is embedded in the school approach to music education.

In this scenario, the research instruments developed for this study can be used to make a comparison between subjective and objective assessments, and will have important insights to offer the teacher. Firstly, by looking at the variable scores which rate students on predetermined learning tasks, the teacher can quickly discover learning gaps. For instance, let us assume that the student is unable to detect if something is out of tune, as would become apparent if the score for Variable 1 (pitch detection) was low. In this case, the teacher could focus on this, before trying to address the problem of the student playing out of tune, which appears in Variable 2 (intonation). Similarly, if the student was unable to hold up the violin correctly, which would be apparent in Variable 4 (pitch angle), corrections to Variable 5 (bow angles) would be diminished, because of excessive bow pressure needed to hold the bow in place when the instrument is slanted downwards (or upwards), along with accompanying orthogonal bow divergence.

Another example, also explored by Reifinger (2009) relates to differences between perception and performance. A component of Variable 8 (reading note durations) examines the participant's ability to read note values correctly. If the student is unable to execute the rhythmic tasks evidenced in Variable 3 (rhythm accuracy), then the accurate reading of note durations will be skewed, possibly because of issues relating to comprehension, rather than implementation. In these cases, the variable scores offer an effective customised diagnostic tool which can help to pinpoint where the learning gaps exist for each individual. This bolsters the authority of the generalists who may be compromised musically, and helps them to act effectively by identifying the best intervention strategy to take.

Developmentally based milestones in string learning are helpful in the cultivation of standards. A process has been put forward to test attainment in these markers of ability, and a correlation with grade outcomes has been established. Furthermore, the research study has indicated a framework through which the observation methods can be used to predict outcomes. The model democratises insider knowledge and empowers generalists to be more proactive in their engagement with performance music learning in classroom scenarios. This contemporary approach paves the way for future research to explore how emerging micro sensor smartphone technologies can be tailored to create mainstream observational learning platforms for use in musical education.

The findings of the study do not seek to replace subjective assessment. As things stand, quality providers of assessment, such as the ABRSM, are an effective way to evaluate student progress in musical instrument learning, whilst at the same time providing crucially relevant learning materials. However, the observational package (POP) developed in the study, bridges a gap between pedagogical concerns during learning, and the performance product on which subjective evaluations are based. Building on work, which began in earnest by Seashore (1938), but using today's technology, this new knowledge signals a paradigm shift in learning expectations. Whilst variable construct criteria required for objective evaluation will alter slightly from instrument to instrument, the data correlation between subjective and objective methods of assessment for strings is compelling.

By its own admission, the ABRSM evaluates candidates on 'musical outcomes' alone, without technical instrumentation specific considerations being exclusively considered. This study conversely looks only at instrument specific considerations and some comprehension variables, which look at a wider predisposition to a musical awareness. Furthermore, the study approach makes clear that music outcomes relating to the performance product, in terms of interpretation and musical talent, are outside the remit of the observations.

Likely scenarios stemming from this will affect the way instrumental preparations take place prior to examination. Using this new knowledge, it can be envisaged how, for instance, a DIY Kinect Motion Capture Studio, as described by Remington (2018), is integrated with Melodine pitch discrimination software, described by Hoenig et al. (2018), and Kinova motion analysis software, described by Guzmán-Valdivia et al. (2013), to combine in a teaching practice that not only provides precision measurement about player progress during learning, but also offers meaningful insight into the effectiveness of instruction methods being applied in real time.

1.12 School, Children, Dynamics

The International Music Council web site recognises the right to learn musical languages and skills and the right to participate, create, listen and be informed about music. These goals are at the heart of music education and the school where the research took place. The children at the school have come to expect music activities as part of their daily routine, and Tuesday mornings at the school sees each child bringing in their sized violin for group string playing class with a music specialist.

Prior to the research data gathering phase, the participating children (both boys and girls) had completed participation and consent forms giving them permission to take part in the research. The school children arrived in the music room in groups of two, in accordance with child protection protocol, for an average of ten minutes on five consecutive Tuesdays in January 2016. The purpose of five observation sessions was to gather data which would be used to populate the ten variable constructs on which the observation comparison with grade outcomes would be made.

The observational sessions focused each time on a singularity which indirectly relates to playing ability. The separate scores on the different points of focus combine to form an instrument specific meta-analysis tool. This approach minimises errors, which can occur in studies which are isolated from conceptually similar lines of enquiry. In other words, several related observation tests can be more informative about participant ability, than a single grade examination. Indeed, exemplary preparation for performance examinations can mask imbedded inabilities, as the high levels of specific preparation can obscure underlying gaps in learning. This topic will be returned to in the literature review section, which looks at work by Sloboda (1984, 1985) in this regard.

The tests were conducted on the bases of the student being satisfied with their own response. For instance, if the student wished to retake any aspect of any of the tests, with the exception of the sight reading test, that was considered perfectly acceptable, with the overall approach being on the basis of the 'best of' rather than just 'one shot at' answering a question or giving a response to any aspect of participant input.
When both children present had completed their participation responses, the next two children were admitted, until all 60 children from 5th and 6th class had completed the

tasks of the observation constructs. Many of the participants did not complete grade examinations, and could therefore not be included in the sample.

1.13 Ethics

The 37 students who participated fully in the study, both school and private, ranged in age from 8 to 18 years. Their identity was coded, and anonymity established throughout the data collection process. The data has been stored securely on a hard drive of a password protected computer, stored in a secure office, and will be destroyed five years after completion of the research project, in accordance with the University of Lincoln's conditions of ethical approval and data protection protocol and guidelines. Participant information sheets were signed and returned by both the participants and their primary caregivers. These documents sought approval for student participation in the observation process. Assurances were given to the students that observations were for assessing learning indicators, rather than individuals. Reassurances were also given that there were no adverse consequences for students who displayed limited musical abilities. It was also pointed out that the research findings had no connection with school records of students while at school, nor did it have any function in relation to competitive attainment goals.

The degree to which views 'should be explained' to the participants is of ethical concern according to Walford (2001, 136). Temperament, for instance, is a culturally imbedded phenomenon and it may be unwise to question indigenous value judgments in relation to tonality. However, my objectivity in these matters is based on pedagogical reasoning discussed in chapter 3, and should stand up to criticism of the approach masking any unjust ethical overtones, like ones found in colonialist or religious discourse. In terms of ethics, the researcher in this case is not trying to defend a view of what music is, but rather make use of structures which reside in the classical music tradition because of their *pedagogical* significance. Students were given the option to withdraw from the observation process at any stage, but efforts were made to share the reasoning behind the observation process with the participants at all stages of the research process.

Whilst established ethical guidelines have been followed, the risks to participants can never fully be addressed according to (Oliver (2004). Before the pilot was undertaken, efforts were made to eliminate participants with difficulties such as hearing loss or amusia – an inability to process musical sounds. Assurances were also given to minimise potential risks, such as students feeling they were in competition with each other. However, there is a competitive edge to any form of assessment or judgmental process, as the students want to do well. It is reiterated, however, that it is not the students who are being judged in this instance, but rather the efficacy of string pedagogy processes and the existence of stages of learning. A case is made for similar early interventions and comprehensive music education and 'intonation opportunities' by Jaccard (2014) citing Willems (2012) who, when talking about 'intratonal space', discusses how the ear becomes more refined, causing the distance between intervals to appear to become further apart, making them easier to recognise. Cognisance of this has helped to make the participants feel more comfortable with what might be viewed elsewhere as challenging scrutiny.

Furthermore, to minimise the possibility of students feeling that they were being judged in a competitive way, reassurances were given through the language used in the student consent form, explicitly setting out what was required of them. The tasks which went to make up the operational definitions consisted of single performance routines which could be repeated. It was pointed out, during the observational process, how the elements being observed reflect teaching, as much as learning. Language was used, such as: "To the best of your ability, can you play for me…" or "You, the participant, can decide which version of a task should be considered in the observation", meaning which 'take' best represented the participant's ability. In other words, a sense of ownership of the process taking place has been given to the pupil. This open-ended, dual responsibility for the efficacy of the learning patterns being studied was considered best practice, to minimise any negative ethical impact.

Instances where this 'partnership' approach would have been substituted by what Hash (2011), describes as a less flexible didactic approach suggest how over-reliance on structured class procedures, dating back to the Universal Teacher method developed by Maddy & Giddins in 1923 for teaching strings noted in Bates (2011), would be unlikely to satisfy ethical requirements today, because of the inappropriate and condescending language used in that approach at the time. Western art music itself, as a pedagogical method, is nowadays brought into question, because of perceived hegemonic origins further complicating the issue.

Advocates of traditional music would perhaps be slow to put intonation forward as a fundamental tenet, owing to ambiguities around regional temperament preferences. Indian music, as a foundation, would require the navigation of twenty-four microtones between each interval, much like ornamentation in traditional music. Jazz musicians are less inclined to favour an approach to music which positions literacy at the centre, owing to their preference for improvised content. The theme on which improvisation is based, however, is ironically faithfully reproduced to the letter. The Universal Teacher method in Bates (2011) retained a language of division that today would be considered to be totally unacceptable ethically, as a model going forward. The condescending language used, for instance, is an illustration of what Born and Hesmondhalgh (2000) was grappling with a century later, in her critical acuity of the deployment of power in Western music. The intonation and bowing components in the study operationalise tonality, equal temperament and tone production as a point of agreement and consensus. Kodály opposed dilutions and substitutions, pointing to the need for intonation teaching practices to begin as early as kindergarten level, as discussed by Jaccard (2014).

The bona fide status of the observational documentation: participation sheets, questionnaires and ethical approval forms, is made clear by highlighting the study's association with Lincoln University. Participants, who are the subject of the observation, should feel confident about the authenticity of the research and comfortable in the knowledge that strict ethical guidelines have been followed through. A copy of the consent form is given in Appendix 1.

Ethical guidelines for educational research concerning a child's ability to be understood in BERA (2011, 6-7) have been considered through the provision of a child-friendly sheet (in addition to one for the carers giving consent) which gives the child the opportunity to agree to participate and prepare for the study. Furthermore, my background and socialisation in music has enabled me to anticipate, with some degree of accuracy, any unintentional duress caused to students. In addition, efforts were made to match ability to task, create a uniformity of design, ensure a transferability of method to other instruments and minimise interference with class work, with observation sessions being kept separate from regular class lessons.

1.14 Research in this area

Former approaches, including: Seashore (1938), Wing (1962), Zenatti (1969), Imberty (1969), Shuter-Dyson (1968) and Gardner (1973), are limited in comparison to modern testing methods. Sloboda (1984) disputed Seashore's pitch discrimination test because of unrealistic assertions about participants' abilities to detect fractions of a semitone. Indeed, Sloboda (1985) uses the term "generous minimum" in making useful determinations about novice string players' abilities to regulate pitch and rhythm. However, Patterson (1974) disputes volume discrimination tests, and points to technical deficiencies in the Wing (1948) and Shuter-Dyson (1968,) tests which make them outdated for contemporary use. Test designers, including: Lowery (1926), Ortmann (1926), Kwalwasser and Dykema (1930), Drake (1933), Lundin (1949), Whistler and Thorpe (1950), Gordon (1965), and Gaston (1968) have narrowed their observational concerns to pitch, rhythm, duration, tempo, memory, intensity, consonance, aesthetics, knowledge, motivation, intervals and transposition. For a comprehensive discussion on former approaches, see Farnsworth (1969).

1.15 Summary

An overview of the text and an outline of what the research consists of has been given. The purpose and reasons why the research has been undertaken have also been outlined. My positionality, limitations and research questions underpinning the research have also been discussed. The hypothesis has been defined and a strategy for testing it has been put forward. Aspirations for the study have been stated and technological innovations have been signposted. The study has been placed in context, with standardisation and clarity being muted as important factors providing impetus and motivation. The research structure and process have been outlined and its limitations have been anticipated. The following chapter conducts a literature review of work done in this area and discusses key points raised.

2 Chapter Two Literature Review

2.1 Introduction

The literature review focuses on topics relating to string learning and assessment, theory and educational philosophy. Desk-based research in this area examined articles from journals including: *American String Teacher, Bulletin for the Council of Research in Music Education, Dialogue in Instrumental Music, International Journal of Music Education, Journal of New Music Research, Journal of Research in Music Education, Journal of String Research, Medical Problems of Performing Artists, Music Education Research, Music Perception, Psychology of Music, Research Studies in Music Education,* and *The String Research Journal* amongst other relevant sources. An overview of the relevant texts informs the hypothesis statement, which postulates a link between detectable modalities of string learning and subjective grade musical outcomes. The inference of the correlations is that scores in particular tasks at the learning stage will have a bearing on performance grade outcomes. Innovation and sophistication in the way assessment is conducted has increased, both through insight into learning processes, and through creative thinking on the part of the research and technology communities. The literature review looked at the following areas.

1 String performance assessment
2 Frequency change detection and intonation
3 Bowing, posture and tone production
4 Music literacy acquisition
5 Emerging technology
6 ABRSM
7 The use of rubrics in musical performance assessment
8 Theoretical framework and educational Bildung

2.2 String performance assessment

Discoveries since performance measurement began in earnest with Small (1937) and Seashore (1960) show how techniques have advanced, warranting a reappraisal of what a parametric measurement, through observation, is capable of telling us. Seashore's Measurement of Musical Talent first appeared in 1915 (Seashore, 1915) and measured pitch discriminations, loudness determinations, rhythm comparisons, note durations, timbre differences and tonal memory. The resultant scores presented a pattern of musical ability similar to standardised tests. Standardised Tests of Music Intelligence in Wing (1962) was more concerned with general modes of musical expression, looking at preferences, aptitude in analysis, phrasing and familiarity with conventions. The Musical Aptitude Profile Manual by Gordon (1965) offered a forced choice decision options design relating to musical sensitivity, tonal and rhythmic imagery. Measures of Musical Abilities by Bentley (1966) again focusing on pitch discrimination, tonal recall, analysis of chords and general expression was particularly suited to class assessment.

These models vary in complexity, but all share similar interests in what Bentley (1966) described as elements necessary for the progress of music. They form the basis of much of what constitutes music assessment today, and have a sound basis in music learning theory. The current research builds on this paradigm and develops aspects that can be exploited with modern technologies.

Learning how to play a violin requires a commitment to continuous learning, and maintaining a standard will require a sustained commitment on the part of the player. Often, achievements are incremental and difficult to perceive. Yet most of the focus in assessment rubrics and performance rating presupposes that a 'final destination' in performance has been arrived at. Most novice learners will not go on to become virtuoso performers, however. Conventional models therefore may be a poor fit for measuring learning that is taking place at an elementary or novice learning stage. A model which is more concerned with whether the student is cultivating the tools needed, to carry out this lifelong process of continuous improvement, may be a more appropriate fit, given what rating scales are trying to measure (i.e. student ability areas of string learning).

It is from this perspective that the research project is conceived. Rather than observing to find qualities in a good performance, the observer looks to see if the tools necessary to achieve a good performance are present or are being obstructed. Teaching methods to practice, rather than teaching methods to play, represents a similar approach. A preoccupation with refinement, aesthetics and artistry are somewhat misplaced at a junior level, as this stage of learning should be more concerned with separating out individual strands of ability, like tone production or sight reading ability, and measuring progress in these areas in a more exacting way.

The segmentation of music learning into its component parts may be resisted, with Wrigley (2011) for example suggesting that the quantification of musical performance is artificial and unmusical. This is, however, what musicians themselves must do all the time when working through repertoire during a private lesson. Separating out each individual nuance is characteristic of the master class lesson scenario also, where specialisation is the thing being cultivated by both master and student. Segmentation therefore is neither unmusical nor artificial, in the context of how music is created.

According to Wind and Engelhard (2015), examiners uphold different constructs. The differing evolutions of the various instrument families are equally valid. Cultural expectations that are associated with each section of the orchestra, for instance, with its own conventions, make a consensual framework which works for all disciplines equally, more difficult to construct. Furthermore, examiners are not schooled in every discipline in which they must examine, with a large burden being placed particularly on piano-playing examiners to excavate for musicianship in areas that are quite different from their own discipline. In addition, bias according to the evaluator's own musical experience and different criteria being employed for differing levels of ability have been noted in Thompson (2003).

Secondary to core constructs of tone, rhythm and posture, which are generally shared by all disciplines, are musical interpretation and understanding. Also known as musicality, these give rise to a second set of constructs, which cross over more transparently between disciplines, and include constructs such as confidence, style and character. The PERS (Performance Examination Rating Scale) in Wrigley (2011) provides a relatively precise diagnostic tool across five instrument categories, and goes some way to accommodate accountability imperatives being placed on raters. The scale provides important feedback for learners about exactly what is required of them, by providing clear definitions of performance goal objectives. Assessment rubrics appear to have improved pedagogical utility according to Latimer et al. (2010).

An accommodation of shared meanings through 'intersubjective' objectivity may be a more appropriate way to measure performance. Disciplinary objectivity, using criteria from a shared code of constructs, is favoured when forming judgments about achievement in artistic performance according to Thompson (2003). Assessment instruments used to measure attainment vary, depending on the family of musical instruments concerned. Measurement instruments, such as the ones used for strings by Reed (1990), brass by Bergee, (1988), selected woodwind by Abeles, (1973), and choral by Morgan (1981) all produce comparatively uneven results. While Whybrew (1971) suggests validity and reliability to be the main concerns in music assessment, Fiske (1983) found reliability was shown to be poor, in instances where performances were repeated by participants, without the prior knowledge of the raters adjudicating them. In other words, examiners, on hearing the same piece of music by the same player without knowing it was the same performance, were awarding different marks.

Despite its popularity, rater performance assessment is brought into question regarding validity, reliability and fairness, according to Wesolowski (2015). Engelhard (1996) found statistically significant differences in rater accuracy. Wind and Engelhard (2015) found the development of indicators of rating quality can inform score interpretation. It was noted that even experienced musicians can be inadvertently influenced, in the way that they make judgments about musical performances by Elliot (1996). Phillips-Silver et al. (2013) recommends a distinction should be made at least between tonal and rhythm aptitudes.

Different strategies are required to take account of varying evaluative criteria needed to assess distinct types of musical performance according to McPherson (1995). To give an example, sight-reading tests and repertoire performance require two entirely different types of observational skills. The reversal of key characteristics of assessment to foster increased learning is advocated by Sadler (2015) as an alternative to conventional approaches. Ensemble and soloist playing are viewed differently in assessment according to Morgan (1981). Order of appearance has, counter-intuitively, been shown to have a significant effect on ranking, with later appearances in competitions ranking higher than ones taking place early on in the proceedings according to Flores and Ginsburgh (1996).

In relation to the above, we can say that rater evaluations can be inconsistent and vary, depending on an examiner's instrument speciality, background and experience, the examination environment, the test model being applied and examination time, place and schedule. Focus is generally placed on musical outcomes and how the performance is received by the examiner. The variation between examiners can affect the outcome for the participant and no feedback in relation to the learning is usually forthcoming in this model of assessment. A common thread running through assessment instruments, however, includes the following elements: rhythm; intonation; tone production; and reading ability. Mills (1991) found that the breaking down of music performance into components, however, is not done without controversy, given that methods using these criteria are sometimes unable to take account of a good performance which may have had some mistakes, or bad interpretations which are flawless.

Efforts to address this problem are evident in the development of 'criterion statements' for music performance found in Swanwick (1996) which look at an overall impression of a performance without direct regard for individual elements. The statements read like that of the informed critic, and demonstrate a balanced appraisal, through the lens of discrimination and insight. Terms like 'erratic and inconsistent' or 'confident technical mastery' describe what is being observed. This type of observation is nevertheless a subjective opinion, albeit an informed one, which demonstrates discrimination and expertise, underpinned by a cultural context where musical traditions are identified. McPherson and Thompson (1998) advises that consideration should be given to social, personal and cultural influences which affect judgments being made in this way.

Assessment of the student, and assessment of the programme, constitute two separate tasks according to Asmus (1999). The collection and interpretation of information gathered about students, and how this information may be used to improve learning, is central to effective assessment procedures. Criterion-referenced approaches set out expectations prior to commencement. Such benchmarks, which describe standards expected in rubrics, introduce objectivity into the process and make informed educational decisions more feasible. Norm-referenced approaches, where the student is assessed in relation to how other individuals have performed, is therefore more problematic, owing to the lack of criteria present in the structure of the assessment.

Authentic assessment involves the assessment of tasks to be accomplished in real-world situations, rather than simply answering questions. Goolsby (1995) indicates that it differs from portfolio assessment, which is an on-going process of assessing student work, and programme assessment, which is concerned with strengths and weakness of a given teaching programme. Summative assessment looks at the overall effectiveness of an implemented programme of study, and formative assessment steers such programmes along as they develop.

Validity (relating to how effective the assessment instruments are at measuring what they set out to measure) and reliability (when repeating the assessment elsewhere whilst maintaining consistency over time) are two important elements to be maintained. The Rasch model, named after Georg Rasch, is a palette of psychometric tests, used for creating measurement from categorical data by balancing respondents' abilities with task difficulty. It represents a structure which data should exhibit to obtain measurements from the data, according to Wesolowski et al., (2015), using a heuristic organising principle. Kafol et al. (2015) indicates musical objectives which favour psychomotor over cognitive or affective domains provide technical expertise in performance, without addressing formalist understanding.

Heneghan (2002, 104) states, "Formalism may be associated with a focus on the great works of art as exemplars of artistic form suitable for study". Objectives which focus only on psychomotor concerns can overlook participants' understanding of study material. Refined motor skills, writes Gzibovskis and Marnauza (2012) are a significant factor in improving coordination and accuracy, but they are a separate matter from measuring comprehension of material. In the research observation study, special attention has been given to balance tasks with comprehension, and it does not measure participants' interpretations of formal works of art.

Broadly speaking, there are two views of assessment in music performance. One view sees music performance as being outside the realm of parametric determination imbued with cultural, personal, subjective, and regional proclivities which exist outside of measurement criteria. This view sees music as being something to be enjoyed primarily, and places little emphasis on standardisation or homogenisation of abilities to make them measurable parametrically. This view is distrustful of testing generally, whilst at the same time being supportive of a competitive drive to succeed and be recognised within a genre circle. In contrast to this view, formal approaches to music study, borrowing on a European tradition, view testing and standardisation as acceptable norms. This model at the outset appears to be somewhat removed from the emotional connection to the music, and can be misunderstood for its insular detached emotional aesthetic. This condition, however, is particularly suited to the re-creation of artistic works, with a primary goal being to represent the composer's work faithfully without personalising it. This difference is discussed in more detail in work by Bernstein (1975) who put forward a theory of codes to explain this fundamental difference in approach.

Both perspectives function differently within an observational process. One being sceptical, the other compliant. One engaging with material, as if it represented them personally, the other standing back. One wanting to have markers of identity recorded in the observations, the other anonymously playing to the letter of scripted detail. One distancing itself from inference of testing, the other committed to critical appraisal as process.

This duality underlying the cultural context is not without its challenges in research observations. Very often a child who has cultivated spontaneity because of learning music by ear may struggle with reading music. A child who reads music prolifically may be inclined to play in a more mechanical way. A child who learned to play without structure may display a comfortable disposition in performance, while at the same time encountering insurmountable technical challenges. A child well-grounded may appear less at ease initially, but more comfortable with challenges. Striking the balance between both valid approaches was considered during observation.

There is no single package currently in existence which observes all of the assessment criteria being utilised in this study. Moreover, discernment of pitch and metre travel along separate trajectories of cognition. There is general, but not unanimous, agreement that assessment is important in music education. A distinction should be made between assessing modalities of learning and the product of that learning. Scepticism can exist about subjective assessment given that varying judgment criteria and settings are employed according to McPherson and Thompson (1998).

2.3 Frequency Change Detection and Intonation

When Helmholtz (1863) put forward the concept of resonance within objects through sympathetic vibration with external ones, he was paving the way for a line of inquiry into the relationship between frequency and tone that continues to this day. The third German edition, published in 1870, of *On the Sensations of Tone as a Physiological Basis for the Theory of Music* was first translated into English in 1875 by Alexander J. Ellis. Ellis added to the original text the term *cent* on page 41 to refine Helmholtz's interval measurements. He achieved this by dividing an equally tempered semitone into 100 parts to determine degrees above or below a given frequency. Ellis (1876, 446) also added appendix notes on the determination and history of musical pitch, using the same system for calculating cents from existing interval ratios. The cent system of dividing a semitone is still widely used today and features in both the pitch determination Variable 1 where notes in the listening test are adjusted in cent, (for instance, the quavers in question 2 are tuned 30 cent sharp), and intonation Variable 2 where the degree to which notes are out of tune is calculated in cent see Figure 13 on page 88.

Barbour (1952) provided scientific justification for equal temperament on which Western art music is based and, in so doing, brought focus in performance, on exact distances between notes. Duffin (2007) reminds us that, despite the establishment of a flattened perfect fifth, the system would be used to construct categories of scale patterns as used for melody building, and corresponding chord constructions from which composers could then modulate freely. The theory of pitch discrimination forms the basis for the first observation variable, which examines the ability to discriminate between those notes which are in tune with a central tonality, and those which are not. Determination of the accuracy in the creation of these notes relative to a given sound, is of interest in music-making. This ability, in turn, impacts on performance accuracy in each key or tonality. In constructing a variable to measure this aspect, the literature review looked at related work.

Hargraves (1986) notes how the mastery of melodic contours is generally accomplished by the end of the pre-school period. Therefore, the ability of a child to discriminate between different pitches, used to create a melody, is already active. Musical development in later childhood is more reflective of accuracy and representation within the pitch-interval relationship (the tonal system). The ability to distinguish between one pitch and another is central in Western music culture which is based on equal temperament. Equally tempered semitones are accurately separated by 100 cent divisions by Ellis (1876). Each degree of a musical scale is defined by a predetermined frequency in Hertz. The construction of the variable utilises this framework to offset test notes by a predetermined number of cent to operationalise the pitch discrimination observation.

Bentley (1966) found that seven-year-old children can discriminate quarter tone intervals (50 cent). Hair (1975) noted how a task structure can influence the success of pitch discrimination testing and Sergeant and Boyle (1980) indicated how higher levels of discrimination can be achieved with simplified methods of testing. Trehub et al. (1986) found the diatonic structure to emerge in children between the ages of four to six years. Because of ambiguities around intonation theory noted in Barbieri and Mangsen (1991) and the problems of the tempered fifth alluded to in Di Veroli, (1991), the context of a pitch discrimination test must be considered. While Jensen & Neff (1993) found a sequential pattern in the development of auditory abilities, Elbert et al. (1995) noted how the transfer of cortical discrimination of the fingers to the fingerboard of the violin is age-related. Functional mechanisms of internal realisation and comprehension of music – tonal audiation, are supported by Holahan et al. (2000) despite divergent views put forward by Fox (2003) about intonation. Harmonic, melodic, corrective and colouristic tunings described by Kanno (2003) draw attention to contextual issues which influence accuracy in pitch discrimination and the intonation quality which stems from it. Visual or tactile feedback notes Lage et al. (2007) is inferior to perceptive listening technique, as explained by Fischer (2013).

Wallentin et al. (2010) found internal consistency (Cronbach's alpha: 0.87) as a result of using Musical Ear Tests (MET). Surprisingly, Vurma et al. (2011), and Geringer et al. (2015) found timbre was found to affect the ability to detect discrepancies in pitch frequency, with the ear being noticeably more tolerant of voice than trumpet pitch discrepancies. Timbre according to Maezawa et al. (2012) is also related to the fingers used to play notes, with computer models now anticipating the most suitable ones to be used. Thomsen (2012) notes how the ability to sing and hear music internally has been found to improve intonation. Norris (2013) maintains that tonal dissonance detection is significantly weaker in 1st and 2nd graders, compared to 4th–6th graders. According to Fancourt et al. (2013), the perception of pitch change, and the perception of the direction of that change, run along separate paths. Furthermore, Phillips-Silver et al., (2013) discovered with amusia sufferers (individuals who have difficulty recognising and discriminating melody) how the recollection of pitch and metre are not cerebrally linked. Fischer (2013) indicated that sympathetic vibrations and resonances within the violin, and in other instruments, are helpful in guiding good intonation. Buss et al. (2014) describes how memory is also a component in the cultivation of good intonation. Jaccard (2014) suggests ear and voice development should happen simultaneously within 'intratonal space' (between the vocal mechanism and the ear), for intonation to develop. The proportions of the fingers in relation to the instrument has an important bearing on intonation, aside from the pitch discrimination faculty according to Çalgan (2015). Standardisation of the relationship between vibrato and the centrality of pitch indicated in Ho et al. (2015) suggests that tone production takes place in isolation from the context of a tonality. Hutka et al. (2015) confirm that musicians develop enhanced acuity in auditory processing and pitch discrimination faculties, beyond what can be expected in the normal population.

We can summarise from the above literature, several important points that speak directly to the way in which the first two of ten variables developed for the study have been constructed. Children as young as six are already articulating tonal awareness. The environment will affect accuracy of testing, in the same way that a sound booth can make it possible to measure more incremental levels of hearing loss, by screening ambient noise out of the process. There are separate unrelated pathways of processing between pitch determination, pitch change direction and rhythm processing, hence the adaptation of singularities in the observation sessions in order to collect data. Memory is an active component in pitch accuracy, as is cultural capital in music, which will affect accuracy in pitch fault detection, and any subsequent precision in intonation which stems from it. Timbre of the original sound will affect tolerance of inaccuracies heard. While every effort is made to minimise any negative effect of factors, such as mentioned above, it remains outside of the remit of the research to intervene or influence outcomes in any way.

2.4 Bowing, posture and tone production

Several developments have made it possible to be more objective about posture in violin playing and learning. Hodgson (1935) first introduced photography for the purposes of making determinations about posture and bow movement. Of the many violin theorists, Gerle (1991) has produced an authoritative account of bowing theory and practice, creating clarity around bowing movement parameters. He describes many types of bowing technique with an authoritative text on how these may be executed. In addition, Gerle (1983) prescribes a system of finger pattern orientations, which amount to a workable strategy for cultivating good intonation. Both these volumes by Gerle clarify left and right-hand technique criteria and provide ample justification for the cultivation of a posture which the student can build on, as more demanding repertoire is undertaken. Together, they provide a sound rational for good posture to facilitate the complicated task set which each hand must undertake independently and together. For instance, the left hand must be positioned under the fingerboard in such a way as to facilitate changes of position later whilst establishing stable finger distance patterns, and the right hand must cultivate a bow movement which ensures optimum orthogonal bow angle with strings for maximum Helmholtz motion. The cultivation of posture which adheres to these long-term objectives is central in the study.

Holding the violin correctly described by Courvoisier (2006), coupled with a sustainable technique demonstrated in Fleisch (1939), can prevent physical problems occurring later on in a young musician’s development. Postural flaws are the result of inadequate technique and are, therefore, avoidable. Araújo et al., (2009) used frontal view recordings to determine postural flaws during performance, demonstrating objectivity in making the determinations.

Posture and bowing parameter measurement have evolved through technological innovation to a high level. The 3d augmented mirror researched by Ng et al. (2007) features technology dedicated to observing the angles of bow trajectories and displaying the results. This important development signalled a new generation of feedback systems designed to help address actual problems encountered by students learning to play the violin.

Schoonderwaldt (2009a) and Schoonderwaldt (2009b) provides considerable insight into bow parameters, and looks at freedom and constraint in subtle bow movements of expert players, demonstrating how subtle nuances can be modelled and monitored. Such systems require considerable lab support, however, and are not suitable for classroom field studies. Wearable technologies, which measure transversal velocity and bow force described in Maestre and Ramírez (2010) also demand considerable physical support. Innovative systems which use technology to support string learning are becoming more portable and less cumbersome and intrusive.

The MusicJacket developed by van der Linden et al. (2011a) is an example of technology being used to assist players to adopt correct posture and bowing technique in real world situations. The aspects being addressed by the device include holding up the violin (variable 4), cultivating a straight bowing action (variable 5) and, when calibrated alternatively van der Linden et al. (2011b) found that the device can also help increase the amount of bow being used by the player (variable 6). This system, which uses vibro-tactile feedback, opens up a new dialogue between teacher and student, creating shared terms to express what is taking place, and cultivates greater body awareness during and after lessons.

Rasamimanana (2012) looked at embodied interaction and modelling of gestures made by musicians while playing, and searched for data-independent analysis tools. By modelling expressive gestures made by musicians during performance with a conceptual framework of 'space of possibilities' (SoP), it may be possible to determine the presence of technique and cognition which they embody. Another study, which extracted variables relating to posture during performance, using a Posture Observation Instrument (POI) was undertaken by Blanco-Piñeiro et al. (2015, 566) and determined aspects, such as maintaining the spine along its 'axis of gravity', the ability of the arms to move freely, and a postural stability frame. These findings show that more attention should be paid to teaching strategies which improve understanding of posture during learning and which then translate into both ergonomic health benefits and improved musicianship.

2.5 Music literacy acquisition

The purpose of focusing on music literacy in the literature review was to understand expected parameters and current norms, which would inform the design of the three variables relating to literacy, which are recognition of note pitch, metric duration, and bow direction indications. Through literacy, permanently recorded versions of past performances, set the past apart from the present, making critical reflection and scepticism a real possibility notes Goody (1963). Research in music literacy sheds light on the following insights and strategies about many of the dynamics taking place within the music literacy learning process.

Mitchell and Green, (1978) point out that the degree to which we understand what we are reading about, will affect our ability to read the material. Subdivisions, complexity, the role of accents, age and experience also have a bearing on musical reading ability according to Upitis (1987) and Drake (1993). Smith et al. (1994) suggest that figural rather than metric representation of rhythmic patterns assists with the reproduction of rhythm sequences, and Waters (1998) found that experts use more eye fixations than novices when sight-reading. This is supported by Reifinger (2006), who states that greater predilection to rhythmic ability results from early exposure. Repp (2006) maintains that phase correction and period correction are central to sensory motor synchronisation, in the coordination and perception of rhythm indications.

Musical independence, through literacy, is a key objective for music educators notes Orman et al. (2007). It is cultivated where literacy experiences are mixed with meaningful music-making activities according to Wiggins (2007). Kopiez and Lee, (2008) advise that three characteristics are essential to good sight-reading: pattern recognition; prediction skills; and auditory representation. Furthermore, Darrow et al. (2009) states that these tasks should be mastered by the second grade, as the efficacy of doing so later is questionable. Maturation, acculturation and active learning are also put forward by Reifinger (2006) as crucial to literacy cultivation. Hayward and

Gromco (2009) found auditory, visual, spatial and kinaesthetic components working together to produce speed and accuracy. A relationship between playing in different combinations of ensemble, and improved sight-reading levels was noted by Dumas (2010). In contrast to Mitchell and Green (1978), Oare and Bernstorf (2010) remind us that sight-reading should not be confused with inadvertently playing by ear whilst reading. Waller (2010) draws attention to the link between writing music and music literacy.

Gudmundsdottir (2010) constructed a system to measure frequency, continuity, and complexity of sight-reading errors, finding them to be largely age related. Rhythm forms of Hayden, popular at the beginning of the 20th century, have been replaced by Kodály methods, to improve developmental music literacy strategies according to Gerber (2011). Sound before sight in Jacobi (2012) as an approach, is reiterated as a way of improving literacy in Musco (2011). Alexander and Henry (2012) rationalise that the key chosen in a sight-reading test, will have an impact on the readability of the music by novice string players, with keys like G and D Major being favoured over flat keys. This is because the finger pattern of the left hand, used to play in these keys requiring the 2nd and 3rd fingers to be close together, is generally introduced before others.

Benedict (2012) argues that a scripted curriculum is an imposition of meaning, suggesting broader constructions of literacy to be more helpful. This is supported by Hansen and Miligan (2012), who state that literacy is a challenge to 'genuine' music learning experiences. High stakes testing places further pressure on conventional literacy in the arts notes Slater et al. (2014). Despite this, Abrahams (2015) suggests a new type of literacy is emerging, given the way young people interact with technology and music learning.

It is evident that different strategies, working in concert, could improve the overall effectiveness of literacy in music. For instance, Allen and Duke (2013) found a link between performance quality and overnight memory precipitation. This is a form of mediation rather than representation. Kanno (2007) also put forward another form of mediation through prescriptive notation, unlocking new ways of understanding notation. Dougan (2015) notes how displacement of conventional ways of engaging with repertoire further alienates students from music literacy programmes. Tomlinson (2015) and Burton (2015) call for a re-evaluation of the role of literacy in music instruction. Grout (2016) defends the central role of literacy as a foundation in music learning. Fehr (2014) indicates the need to instil this through standards, assessment and evaluation. Altenmüller et al. (1997) remind us how the re-creation of musical ideas is dependent on reading symbols.

Strategies worth noting in the literature, which inform music literacy assessment, include Jentzsch et al. (2014) who observed how a reduced corrective response in advanced players strikes a better balance between stopping and starting during performance, and accuracy in reading. Mishra (2014) found the ability to predict what is coming next directly affects a participant's account of a piece of music. Grachten and Krebs (2014) brought attention to how an understanding of the placing of score dynamics can also be a factor. A link between auditory processing and literacy is made by Steinbrink et al. (2014). Other strategies, including score analysis and mental practice in Fine et al. (2015), and sight singing in Sheridan (2015) enhance a 'continuum' of aids and contribute to better preparation and good performance.

The review indicates how continuum in music education is better served with literacy strategies. Gwen Moore (2012, 63) noted "positive affirmation of those from classical backgrounds but left students from other musical backgrounds doubting their musical ability, in particular vis-à-vis their theoretical/technical skills." Moore further asserts that "induction practices in higher music education for students of diverse musical backgrounds requires consideration" (Moore, 2012, 75). Heneghan (2002, 16) notes "the corpus of knowledge generated by western art is too valuable a resource to be squandered." Gamble (1988, 26) states, "Elementary education majors should be required to complete coursework in music education to make them musically literate… and to help them realise that music is an essential part of their own education and…students in their future classrooms."

Aguilar and Richerme (2016) found music teacher educators to favour National Music Standards, over Race to the Top incentives, and STEAM approaches in music education. This view is supported by Frederickson (2010), who suggested that standards guide teachers in creating objectives that are necessary for music students to know about. Furthermore, Cangro (2016) reiterates, standards can be raised by students taking responsibility for their artistic development through collaboratively working together. Elmore (2006) points out that a system of accountability for students and teachers can be achieved with standards-based, data-driven instruction, suggesting that summative assessment simply reflects success or failure of a system.

Practices surrounding music literacy acquisition such as enculturation, understanding of the material, memory training, early exposure, eye fixations, along with others have been explored. There appeared to be a gap in the literature in relation to reading and assessment of bow indications.

2.6 Emerging Technology

There is a long history of technology in music education. Since the invention of the metronome in 1815, or the phonograph in 1877, steady advances have been made to enhance learning using technology. Indicators of ability have been the subject of interest to musicologists and music education researchers throughout the last century. The inquiry, however, has not kept pace with exponential developments in technological measurement techniques.

Seashore's scientific account of measurable parameters in string playing illustrates how highly personalised traits – such as vibrato, intensity and intonation – were profiled in order to compare them to other string players. Attributes examined for musical significance included frequency, intensity, duration and form, corresponding to pitch, loudness, time and timbre. Indeed, anticipating changes in technology, Seashore (1938, 30) foresaw that: "In the future, musical aesthetics will be built upon the basis of scientific measurement and experimental analysis." Moreover, his approach assumes that "quantitative measurement of performance may be expressed in terms of adherence to the fixed or so called 'true', or deviation from it in each of the four groups of musical attributes." (Seashore, 1938, 30).

Drawing on the findings of Small (1937), many of the assertions made by Seashore in defence of his measurements would not stand up to robust appraisal today. Deviations of 1-5% of a semitone, for instance, are not generally audible, as they pertain to a bandwidth of vibrato, rather than intonation. The approach does, however, illustrate the potential for objectivity in performance analysis and a perspective from which the study at hand is situated. Technological advances in the intervening years afford current research greater scrutiny and interrogation of observational parameters and their analysis.

Technology has made it possible to track posture and bowing, and make determinations about pitch and intonation in an exacting way. Rating scales such as Gordon (2002) provide a way of evaluating achievements, that are criterion specific. Music educators are being increasingly compelled to gather, document and track data relating to student ability, see for instance Wesolowski (2015). Dittmar et al. (2012) encourages music education stakeholders to embrace these new developments and reiterates the need for the music and technology communities to work more closely together, to reach a technologically literate generation of emerging musicians.

This growing interest in technology and music is reflected in the increased number of Music Education National Conference MENC sessions dedicated to it notes Palkki et al. (2016) where these new technologies are discussed. In performance measurement described by Schneider (2015), it is necessary to ensure that the technique of measuring does not interfere with what is being measured notes Metcalf et al. (2014). Exponential improvements in accuracy and verification are taking place giving greater reliability, through transparency and repeatability.

The design and methodology of the study, however, strives to minimise intrusion posed by technology on participant involvement, in favour of methods which ensured freedom of movement, participant spontaneity and universal application in the field. It was felt that delicate motor responses, needed on the part of the participant, would be compromised by such intrusions, leading to verifiability issues when replicating the observations elsewhere. To avoid this problem, and still achieve the high level of accuracy needed to make objective determinations in the 10 variable constructs, audio recognition software was deployed.

It was decided to approach the problem of data collection relating to participants' responses from an audio and visual analysis perspective alone. This would minimise any changes to normal performance conditions, while maximising the 'depth' of analysis of sounds being made. Innovations in the area of audio analysis have made it possible not only to determine note accuracy, based on audio analysis of notes played in real time described byDietze (2013), but also to separate out complex polyphonic textures post-recording in keeping with a 'sound alone' analysis put forward by Parncutt and Mc Pherson (2002).

In effect, the quality of analysis of audio material has increased, despite the volume of data needed to make the determinations being reduced. This is because each new sample becomes a verification of a characteristic in the playing, rather than a new discovery. For instance, relating to intonation, it can be shown, through a detailed visual profile of audio events particular to a given observation, not only to what extent a note is out of tune, but exactly where it occurs. This is because the same characteristics of intonation defect tend to be manifest in scales, pieces and sight-reading examples. The study of sound waves is where such an analysis can be made possible according to Seashore (1938). In a similar way to the simplicity of approach to analysis of sounds, observations of elbow angles at key points, affect bow trajectory and determine the quality of the sound , see Guzmán-Valdivia et al. (2013).

New technologies described in Dittmar et al. (2012) are bridging the gap between music education and computer science are being developed all the time Music Information Retrieval (MIR) constitutes a branch of computer science connecting technology and education. Some of these developments including the Interactive Music Tuition System (IMUTUS), the Virtual European Music School (VEMUS), and the Interactive Multimedia Environment for Technology Enhanced Music Education (i-Maestro), are part-funded by the European Commission. The Music Representation Research Group (IRCAM) in Paris has developed a score following system (Antescofo), used for interactive accompaniment. Additionally, Tonara, an interactive sheet music app, automatically detects the user's position in the score in real time, from the player's microphone input.

It is important, however, to link the technology to pedagogical objectives rather than let it dictate them. For instance, measuring by how many incremental degrees the violinist holds up the violin is not as helpful in learning as is making a general determination about the students' tendency to hold up the violin. Blanco-Piñeiro et al. (2015) included, as a category in musician posture observation, the idea of making determinations based on stills of the playing pose adapted by the participant, without them actually playing. This can be useful as posture captured as a pose can be as revealing as the actual playing, given the changeability of a performer's movements during performance.

Another study by Charles (2010) identified crunching, skating, player nervousness, intonation, bow bouncing, extra note, sudden end to note, poor start to note and poor finish to notes as characteristics of novice players by using Music Information Retrieval (MIR) analyses of waveform signals alone, to detect the presence of these faults. Results from this study indicated how, when compared to baseline characteristics of advanced players, it is possible for a computer to differentiate between two levels of player and detect playing faults, and found that beginner samples contain more power in the unwanted frequency ranges than professionals' ones. When more than one feature was used to represent the data, 97% accuracy of the test was achieved. Furthermore, it was shown that performance inaccuracies do not occur in isolation.

A summary of literature covering technologies which have been used to retrieve information from the bow is detailed in Pardue et al. (2015). The bow tracking system using near-field optical reflectance sensors was used to achieve bow position determinations on the string by measuring the triangle's apex angle (the angle between bow end hair extremities and string), with four sensors placed along the bow, to calculate the bow's location. While the sensor technologies do not place requirements on the surrounding space, the attachment of sensors to the bow itself would appear to cause a restriction in the player's movement, the feel of the bow, and consequently, spontaneity in the playing, thus ruling out this option for the study at hand.

Computer-aided platforms can also provide automatic scoring systems and self-learning experiences creating formative and objective assessment tools. Playing mistakes were more easily identified with concatenation of segment-level features, rather than note-level features alone. That is, by observing note configurations rather than observing individual notes themselves. Other approaches include automated music transcriptions for practice purposes described by Zhang and Wang (2009) including frequency alignment for voice assessment, Molina et al. (2013), pitch training systems for violinists Wang et al. (2012), and expressive detection Barbancho et al. (2013) are all further examples of MIR strategies with objective assessment potential. Making use of established MIR tasks described by Tzanetakis and Cook (2002), a study by Wu et al. (2016) comparing audio features and design features in a regression model found general feasibility in the assessment of student performances without reference to a music score, leading them to argue:

> "A model that automatically generates assessments from the audio data would allow for objective assessments and enable musically intelligent computer-assisted practice sessions for students learning an instrument." (Wu et al., 2016, 99)

Regarding formative learning, the potential for unbiased objective evaluation, assisted by music information retrieval systems is evident. Other score-independent approaches described in Nakano et al. (2006), and Mion and De Poli (2008) make comparisons between non-determined and pre-determined audio data. Score-dependent approaches such as Abeßer (2013) model the relationship between score-based features and expert musicians' ratings. Mistakes in performance can be detected with automatic music transcription using 'audio to score alignment' (Fukuda, 2015). The 'Sparse Coding' (structures inherent in data) feature learning method, as described in Abdallah (2006), was advocated as a solution to solving high domain knowledge requirements of music information retrieval systems. Indeed, it is stated that "adaptive sparse coding can discover musically relevant structures in polyphonic mixtures, yielding accurate transcription." (Abdallah, 2006, 192).

We can take as given from this review of collaboration between music education and technology that the relatively new branch of computer science research in the area of music information retrieval for applications in interactive teaching and learning environments, has the potential to revolutionise future approaches to music assessment in an objective direction. An exponential growth in conferences, seminars and journal articles dedicated to this new field is testimony to this development.

Technology linked to the approach at hand includes the i-Maestro project funded by the European Communities Sixth Framework Programme on Technology Enhanced

Learning (IST-026883). This is an innovative approach to delivering feedback about posture in string playing, guided by an understanding of pedagogical needs. It links technology and string learning through the development of a music information retrieval system (MIR) for tracking posture and bow trajectory. The results are projected on a screen for gesture analysis, and formative learning feedback to students, during learning.

Based at the University of Leeds, a VICON 8i optical motion capture system, uses 12 high-speed infra-red cameras to map the 3d space in which the observations are made. A Max MSP/Jitter multimedia programming environment (cycling 74.com) is used to develop the AMIR application. For a detailed description see Ng et al., (2007). At the time of writing, it was unclear if this system or other systems such as Schoonderwaldt & Demoucron (2009) have been adapted to work with more affordable motion capture devices. While the 'MusicJacket' put forward by van der Linden et al. (2011a) provides a vibro-tactile feedback solution to correct posture and bow trajectory during learning in the field, it has not been adapted for observational measurement.

Ways of digitising parameters for correct posture and bow trajectory bandwidths are becoming imbedded in software designed for teaching and learning. It is anticipated that future studies will combine emerging smartphone processing capabilities with MIR systems to achieve greater verifiability, through exponentially increasing sample sizes and greater co-operation between technological and music communities. Currently, there is no "off the shelf" affordable package apart from the one developed for this study, which can be adopted to measure all of the constructs defined in Variables 1 to 10 - the Primary Observation Package (POP) in the field. The technologies which have been utilised during the five observation sessions in the study, including Kinovea for visual analysis and Melodine for audio analysis, are discussed in detail in the section on methodology.

Data streaming and amalgamation systems which talk directly from sensing devices to computer applications and mobile phone apps, are likely to dominate future research in this area. It is worth mentioning that, with greater analysis capability, more can be gleaned from less. This reality is demonstrated by the following extract from an article in the March 2016 edition of *theEngineer*

> "Placing a smartphone next to a printer, the team was able to capture acoustic signals that carry information about the precise movements of the device's nozzle. These signals can then be used to reverse engineer the object being printed" theEngineer (2016)

In a similar way, when applied to music, the analysis of the spectrum of elements present within live and recorded sound make it possible to accrue accurate inferences about the mechanics of how those sounds are created. Conversely, blemishes in optimum tone production can be explained by these elements not being present. In the section on methodology, Figure 13 on page 88 illustrates attention paid to intonation discrepancies, and Figure 18 on page 104 demonstrates how defective orthogonal bow angles (a characteristic problem in string learning) have been measured in the study.

Arising from the literature review, many software options were considered. It was decided that both Kinovea software described by Guzmán-Valdivia et al. (2013) and Melodine software detailed in Dietze (2013) and Neubäcker (2011) would be incorporated into the research design and observation programme, as both of these packages provided workable solutions to the actual observation tasks at hand. Doubtless, more sophisticated methods exist, but the close link with actual teaching objectives has ensured that the involvement of technology has not overshadowed the problems which the research sets out to address.

Technological developments such as music information retrieval (MIR), motion capture and audio analysis are transforming the concept of assessment from a model of informed, insider, summative connoisseur opinion, to one of overt, formative verifiable, objective analysis. These changes present both challenges and opportunities for music assessment design. Challenges, for instance, exist around getting the right mix of technologies to measure what is required to be measured in the classroom or private studio practice. Secondly, the data entry sources must make measurements without impinging on the performer's spontaneity. Thirdly, the design models of assessment should make inquiry about constructs which are formatively beneficial to participant engagement. Measurements and determinations should be easily interpreted, and have a direct bearing on learning. The technologies should avoid being confined to a laboratory or be overly complicated. They should be designed from the point of view of learning, linked directly to a menu of pedagogical objectives and have a direct route back to the site of learning. The design model should be scalable and be calibrated to reflect different levels of ability where possible. The opportunities that this approach presents include low cost formative assessment models for student and teacher, the realisation of objectivity in assessment, verifiability and scalability for further research, comparison with subjective outcomes and transference of ownership of learning to the student.

2.7 ABRSM

The genesis and rationale of the ABRSM, whose assessment system forms the backcloth for this thesis, warrants some consideration. The ABRSM was formed in 1889 following an uneasy truce between the Royal College of Music and the rival Royal Academy of Music, after the latter's examination system (started in 1880) collapsed amidst corruption and fraud. Its genesis was prompted by what Thomson (2013, 111) has described as "the admirable Victorian ethos of self-improvement and its thirst for formal qualifications." However, Salaman (1994, 209f.) argues that "the stated purpose of founding the ABRSM was to raise the standards of performance among applicants for places to the Royal Academy of Music and the Royal College of Music and … The examinations were for the benefit of the colleges, not of the students." Its overwhelming success in providing a musical examination system that operated not only in the UK, but across the British Empire, led to the establishment of similar national music examination boards in both Canada and Australia (Brightwell, 2013). Such was its dominance that, Sloboda argues,

> "its effortless and unchallenged projection of its own rightness about what constituted proper musical activity and learning communicated to me and my peers that British establishment musical sensibilities were globally superior." (Sloboda, 2012, 5)

This dominance led to "increasing criticism of its institutionally inbuilt complacency and conservatism" (Thomson, 2013, 111). However, because most teachers and parents familiar with the ABRSM system continue to see the grade exam format as representing something of the gold standard in instrumental testing, there is little incentive for the Board to undertake a revision of its examination format, with or without the use of new information and communication technologies. The Board's reluctance to change is also a function of its symbiosis of both educational service and commercial enterprise.

In his official history of the Board, Wright (2013, 1) relates that it was "designed to provide for the *objective assessment* of progress in learning an instrument or voice, and applied on an industrial scale" (my emphasis). However, as Wright makes clear, although the Board established a common format for examinations, subjective connoisseurship was more evident than scientific objectivity in the examination process. Indeed: the ABRSM:

> "relied upon the distinction of its examining panel to establish its primacy, and from the outset it was resolute in holding to the line that its examiners' judgement and behaviour were beyond reproach." (Wright, 2013, 82)

Hence the musical expertise of ABRSM examiners guaranteed the reliability of their assessments, which were thereby unquestionable; indeed, it was not until 1993 (more than a century after its foundation), that the ABRSM established a formal complaints and appeal system. However, in 1998 the ABRSM commissioned three academics in the Music Research Group at Leicester University to undertake an examination of the reliability of assessments in the ABRSM. Highly unusually for university research, the ABRSM required the three academics involved to sign a non-disclosure agreement. Although the authors of the report were willing to allow an examination of their report (Hargreaves et al. 1998) for this thesis, a request to the ABRSM to quote in this thesis from the report was denied, despite the fact that the report is now over 20 years old.

However the ABRSM history by Wright (2013) does provide some of the report's findings. Hence Wright reports (p. 232f.) that "examiners in specialist subject areas where they felt less secure (such as singing) tended to play safe, especially in subjects (such as percussion) which they encountered only occasionally", although Wright does not describe how "playing it safe" affected students' grades, but his description is not suggestive of an unbiased approach to grading. Additionally, it was found that "marks for piano tended to be lower than for other instruments, and there was more variability when it came to marks awarded for the higher piano grades" (p. 234). However, in respect to "the Board's policy of using generalist examiners to assess the musicality of the playing rather than its technical achievement," Wright reported that Hargreaves et al. (1998) found that "when these generalist examiners were marking their own instruments, they tended to give lower marks" (p. 234). Hence when examiners marked their own instruments they would assess both the musicality and the technical achievement, and were more critical than when they marked the playing of other instruments, on which they could only assess the musicality.

The report also identified some unusual anomalies for which there were no discernible reasons, for example "those with intermediate service as examiners tended to award lower marks overall than examiners with either shorter or longer periods of service" (p. 243). Similarly, it was found that "female candidates tended to get higher marks than male candidates, and female examiners awarded higher marks." Furthermore Wright also reports that the study found that "marks given for exams involving non-gender-stereotyped instruments tended to be higher than for instruments which had accrued traditional gender stereotypes. But of the stereotyped instruments it appeared that the male ones (percussion, trumpet, guitar) were markedly higher than the female ones (flute, violin, piano)" (p. 234). It is difficult to see what specific factors might account for such anomalies, raising the possibility that they may arise because the subject examination benchmarks established by the ABSRM are insufficiently rigourous or are not applied consistently. According to Wright (2013, 234) the ABRSM's response to the report indicated that the "gender aspect was clearly a complicated issue that needed careful treatment at examiner's seminars, because a bald statement, rather than a nuanced response, could easily produce a cautionary overreaction." Such a defensive response, rather than a spirited rebuttal, plus the disinclination to either repeat such a statistical analysis, or allow the previous report to be cited, after 20 years, does not suggest that the Board has sufficient confidence in the robustness of its grading system to allow it to be submitted to external statistical scrutiny.

Clearly, objectivity, without instrument specific knowledge (as may occur in ABRSM gradings) is incomplete, particularly in an age when technological innovation can gather relevant information on musical performance to validate subjective gradings.

However, the ABRSM examiners have no recourse to such technologies and use non-instrument specific criteria when making an assessment of a performance from Grade 1-8 standard. By examining the musical outcome, rather than the technical means to produce it, the rubric takes on a subjective quality which makes it ideal for the comparison study at hand, which is totally objective in its nature. In ABRSM practical examinations, each element is assessed by subtracting or adding to the required pass mark rather than deduction from a maximum or addition from zero. First, students play three pieces for each of which they are awarded up to 30 marks, with the examiner determining the mark in accordance with their assessment of the student's performance in relation to pitch, time, tone, shape, performance. The elements of pitch, time, tone, mirror the objective interest while shape and performance are the remaining categories broadly deemed to contribute to a 'musical outcome'. Second, the students are then required to play scales and arpeggios, (or unaccompanied song) for which they are awarded a mark out of 21. Third, students take a sight reading test of a musical script, for which they are awarded a maximum of 21 marks. Fourth, students have aural tests for which they can be awarded a maximum of 18 marks. Hence the maximum mark that students can attain is (3 x 30) + (2 x 21) + 18 = 150. The rubric in Figure 2 is taken from the 2018 ABRSM website guidelines and uses subjective terms such as highly, largely, generally, frequently, mainly, suitable, just, unsuitable, some, etc. leaving much room for ambiguity, despite the rubric being connected to a formal marking structure.

All ABRSM examiners have to be piano players to a Grade 8 or higher (not least in order to perform the aural tests), but they do not have to play the instrument that they are examining. Hence it is possible that an examiner, who is a flautist, may be required to assess a pupil's violin playing skills. Hence most examiners are able neither to undertake an instrument specific technical appraisal of a pupil's musical abilities nor to give detailed feedback on their technique. Delivering feedback on a musical outcome only, without instrument specific criteria, places further burdens on the model as the focus must, by necessity, be placed on product rather than process. This assessment limitation, despite being formulated by trained, enculturated and well qualified connoisseurs, lacks a certain specificity which is of particular value to the learner, during and after performance. Bearing in mind the candidate has entered a clearly demarcated instrument category, it is the ability to play the instrument rather than the ability to make music which should be scrutinised. This feedback gap is evidenced where a batch of reports often includes generic recurring themes such as ‘some slips’ or ‘keep it moving’.

PIECES	Pitch	Time	Tone	Shape	Performance
Distinction 27–30	• Highly accurate notes and intonation	• Fluent, with flexibility where appropriate • Rhythmic character well conveyed	• Well projected • Sensitive use of tonal qualities	• Expressive, idiomatic musical shaping and detail	• Assured • Fully committed • Vivid communication of character and style
Merit 24–26	• Largely accurate notes and intonation	• Sustained, effective tempo • Good sense of rhythm	• Mainly controlled and consistent • Good tonal awareness	• Clear musical shaping, well-realised detail	• Positive • Carrying musical conviction • Character and style communicated
Pass 20–23	• Generally correct notes • Sufficiently reliable intonation to maintain tonality	• Suitable tempo • Generally stable pulse • Overall rhythmic accuracy	• Generally reliable • Adequate tonal awareness	• Some realisation of musical shape and/or detail	• Generally secure, prompt recovery from slips • Some musical involvement
Below Pass 17–19	• Frequent note errors • Insufficiently reliable intonation to maintain tonality	• Unsuitable and/or uncontrolled tempo • Irregular pulse • Inaccurate rhythm	• Uneven and/or unreliable • Inadequate tonal awareness	• Musical shape and detail insufficiently conveyed	• Insecure, inadequate recovery from slips • Insufficient musical involvement
13–16	• Largely inaccurate notes and/or intonation	• Erratic tempo and/or pulse	• Serious lack of tonal control	• Musical shape and detail largely unrealised	• Lacking continuity • No musical involvement
10–12	• Highly inaccurate notes and/or intonation	• Incoherent tempo and/or pulse	• No tonal control	• No shape or detail	• Unable to continue for more than a short section
0	• No work offered	• No work offered	• No work offered	• No work offered	• No work offered

Figure 2. ABRSM Marking Criteria.

PERSON SPECIFICATION

ABRSM Examiner: Practical Grades

ABRSM

	Essential Requirements	Essential	Desirable	Evidenced
I.	**EDUCATION & TRAINING**			
1.	Music degree or equivalent musical and technical knowledge and expertise	**Essential**		Application Form
2.	Instrumental or vocal training to professional level	**Essential**		Application Form
II.	**EXPERIENCE, KNOWLEDGE & UNDERSTANDING**			
3.	Music teaching experience (either private, classroom or academic)		**Desirable**	Application Form & Interview
4.	Knowledge of ABRSM's work and ethos, particularly relating to graded music exams	**Essential**		Interview
5.	Experience of acting as an ambassador or representative		**Desirable**	Application Form & Interview
III.	**SKILLS & ABILITIES**			
6.	ABRSM Grade 8 or equivalent standard of piano playing skills	**Essential**		Interview
7.	Ability to mark to criteria and write perceptive, legible and appropriate comments at speed	**Essential**		Interview
8.	Ability to multitask effectively and remain calm under pressure	**Essential**		Application Form & Interview
9.	Flexible approach to a demanding workload and the ability to keep to time	**Essential**		Application Form & Interview
10.	Ability to deliver the Aural Tests accurately and fluently (see Specimen Tests for examples)	**Essential**		Interview
11.	Ability to adapt according to the situation and react appropriately to the unexpected	**Essential**		Interview
12.	Ability to communicate effectively and develop a rapport with people of all ages and from all walks of life	**Essential**		Interview
13.	Ability to concentrate for long periods of time	**Essential**		Application Form & Interview
IV.	**PERSONAL QUALITIES & ATTRIBUTES**			
14.	Respectful and encouraging manner towards candidates	**Essential**		Application Form & Interview
15.	Patience and professionalism irrespective of circumstances or standards of performance	**Essential**		Application Form & Interview
16.	Dedicated and receptive approach to training and continuing Professional Development	**Essential**		Interview
17.	Sensitivity to and awareness of the needs of others	**Essential**		Application Form & Interview
V.	**MISCELLANEOUS**			
18.	Must be available for a minimum of 6 days (5 of which should be Monday-Friday) for examining during each of the three UK exam session each year	**Essential**		Application Form & Interview
19.	Willingness to travel within the UK	**Essential**		Application Form & Interview
20.	Willingness to travel; to international centres		**Desirable**	Interview
21.	A clean, enhanced disclosure from the DBS	**Essential**		Interview
22.	An interest in music education and a desire to support and encourage musical development and achievement		**Desirable**	Application Form & Interview

Figure 3. ABRSM Examiner Criteria

The criteria for examiner selection confirm this, given that in addition to qualifications which would be expected for the role, the piano is stated as being the only essential instrumental requirement for an examiner who may be required to examine string or wind players (See Figure 3). This is itself problematic, given that the role of examiners, as stated in the ABRSM (2018) literature, is "To conduct the highest possible quality assessment of students in practical graded music exams." Many of the candidates expect the examiner to have a level of proficiency on the instrument in which they are being assessed, given the high costs involved in taking practical music examinations. The ABRSM (2018) examiner description material above summarises the examination criteria.

2.8 The Uses of Rubrics

Johnson's analysis (1997, 281) argued that, with respect to music, "criteria in the performance examination cannot guarantee objectivity of assessment, for which we must continue to depend on our examiners' experience and integrity." Johnson's finding that objectivity is unachievable during the examination of music performance, highlight the problem with trying to achieve valid reliable and authentic assessments in music examinations by differing examiners, as occurs with the ABRSM. However, marking rubrics can enable examiners to identify specific standards and achievements. Marking rubrics divide a specified fixed task (in this instance, a musical performance) into its constituent elements (e.g. as has been seen for ABRSM assessments: pitch, time, tone, shape, performance) and present descriptors which correspond with the levels of each element, for example, in the case of scales and arpeggio playing, the highest level in the ABRSM assessment is Distinction 19–21, characterised by: highly accurate notes/pitch; fluent and rhythmic; musically shaped; confident response. The purposes of the rubrics' descriptors are two-fold: first, to provide feedback for the candidates in respect to the different elements of the examination and what they need to do to ameliorate their future grades; second, to determine the marks to be awarded to the candidate.

Following accepted practice, Quinlan (2006, 25) divides rubrics into analytic and holistic. Holistic rubrics evaluate an entire project and yield one numeric score (usually between 0 and 4 or 1 and 6). For example, in a four-point rubric, 4 points may be awarded for exemplary work – typically beyond expectations; 3 points will be awarded for good, solid-quality work (the benchmark); 2 and 1 points are awarded for projects that are below par. Hence a disadvantage of using holistic rubrics is that few details are provided for feedback. By contrast, analytical rubrics consist of multiple scales, and individualised assessment information which thereby provide a more detailed analysis of the performance, such that candidates can see exactly where their strengths and weaknesses lie. In addition, analytic rubrics give teachers the freedom to weight the score of a particular attribute that they wish to specifically emphasise.

The following examples of these two common (holistic and analytic) models of rubrics are taken from Wesolowski (2012). It can be noted that in both rubric examples, latitude is left to the examiner to assess the ability of the candidate. It can be seen, for instance, in the holistic model in Figure 4, which was designed as an individual sight reading test with an ensemble, that terms like "few", "overall", "not consistently" and "moderately" are used to determine ability in a task specific holistic rubric which is scored between 1 and 4. Whilst these terms are accurate descriptors in themselves, they lack an objective quality which is repeatable and verifiable. It is also worth noting that, although this rubric is for a sight reading test, what is actually being assessed are other elements, quite separate from the task of reading music, such as a sense of rhythm, tone, intonation and consistency. As well as a lack of objectivity, another deficiency of such rubrics is that they provide very little information about the student's ability to articulate primary tasks which exist within music literacy, such as reading written note pitches, or the written time demarcation for notes. This deficit makes practical formative feedback ineffective in relation to these primary tasks. Particular aspects, in addition to these two basic elements, can create a finer grained account, such as in the case of string players, the ability to read bow indications accurately.

Example of a Task-Specific, Holistic Rubric Designed as a Quarterly Assessment of Individual Student Sight-Reading within the Ensemble

Quarterly Individual Sight-Reading Performance Assessment	
4 points (Accomplished) ***Grade: A*** *Intonation and balance:* Student plays with good overall intonation and demonstrates ability to match pitch quickly and fluently. *Fluency of reading notation in ensemble context:* Reading is fluid with very few to no pauses. *Phrasing:* Interpretation of melodic contour, note emphasis, and note duration are performed consistently in a proper, idiomatic style.	
3 points (Proficient) ***Grade: B*** *Intonation:* Student plays with moderately-good overall intonation and demonstrates some difficulty matching pitch. *Fluency of reading notation in ensemble context:* Reading is moderately fluid with occasional pauses. *Phrasing:* Interpretation of melodic contour, note emphasis, and note duration are performed somewhat consistently in a proper, idiomatic style.	
2 points (Developing) ***Grade: C*** *Intonation:* Student does not consistently play with good overall intonation and demonstrates difficulty in matching pitch. *Fluency of reading notation in ensemble context:* Reading lacks fluidity with frequent pauses with little coherency. *Phrasing:* Interpretation of melodic contour, note emphasis, and note duration are performed mostly inconsistent of the idiomatic style.	
1 point (Beginning) ***Grade: D*** *Intonation:* Student does not play with good intonation and demonstrates little to no awareness in matching intonation. *Fluency of reading notation in ensemble context:* Reading lacks fluidity and is disjunct with no coherency. *Phrasing:* Interpretation of melodic contour, note emphasis, and note duration are performed inconsistently in the idiomatic style.	
Comments: *You have come a long way! Continue to work on your intonation exercises with a tuner, particularly in the upper register! It continues to be a bit sharp at times.*	Score: *3*

Figure 4. Holistic Rubric Wesolowski (2012, 39)

Figure 5 shows the analytic rubric example produced by Wesolowski (2012). As with the previous rubric, relative terms like “little”, “some”, and “overall”, appear in the determinations check boxes. While it can be argued that these descriptors are capable of some selective objectivity, it is clear that the expertise of the examiner, particularly on the instrument concerned, will contribute to the accuracy of the test model. Being mindful of these shortcomings, Wesolowksi et al. have rightly argued that:

‘As music organizations such as the National Association for Music Education (NAfME) and the Associated Board of the Royal Schools of Music (ABRSM) move towards assessment models that demand a need to standardize and benchmark music performance assessments, holistic scoring procedures are not suitable. Therefore, the demand for reliable, valid, and equitable trait-specific scoring mechanisms is increasing for performance-related assessments in music’ (Wesolowski et al. 2015, 148).

Example of a Task-Specific, Analytic Rubric Created as a Weekly, Individual Music-Preparation Assessment for an Advanced High School Wind Ensemble

Weekly Individual Music Preparation Assessment					
Learning Outcome	**Beginning**	**Developing**	**Accomplished**	**Exemplary**	Score
Rhythmic fluidity within the melodic line	☐ 1 pt. Student plays with little or no fluidity.	☐ 2 pts. Student plays with some fluidity, but does not improve with coaching.	☑ 3 pts. Student plays with some fluidity and improves with coaching.	☐ 4 pts. Student plays with fluidity and self-adjusts.	*3*
Tone control within varying registers	☐ 1 pt. Student plays with little or no control.	☐ 2 pts. Student plays with some control, but has not shown improvement.	☐ 3 pts. Student plays with some control and shows improvement with coaching.	☑ 4 pts. Student plays with control and self-adjusts.	*4*
Control of intonation and demonstrated intonation adjustment	☐ 1 pt. Student needs considerable attention to intonation problems.	☐ 2 pts. Student has some problems with intonation, but does not improve with coaching.	☐ 3 pts. Student has some problems with intonation and shows improvement with coaching.	☑ 4 pts. Student plays in tune and self-adjusts.	*4*
Consistency of focus in rehearsal setting	☐ 1 pt. Overall performance is almost always inconsistent.	☐ 2 pts. Overall performance is generally inconsistent.	3 pts. Overall performance is generally consistent.	☐ 4 pts. Overall performance is consistent.	*3*
Comments: *Continue to work on your fluidity in performance: the connecting and musical shaping of phrases, evenness of notes (in rapid passages), and creating a relaxed feeling during technical passages so as not to rush. Overall, you are gaining much independence in making musical decisions and adjustments. Continue working hard! Bravo!* Sum score: *14—A*					
Explanation of Scoring: Each proficiency level is scored according to the label: Exemplary (4 points), Accomplished (3 points), Developing (2 points), and Beginning (1 point). Students are graded according to each criterion under the level and the total should be added. The maximum score for this rubric is 16 points. The scoring cut points are as follows: A: 14–16 B: 11–13 C: 8–10 D: 4–7					

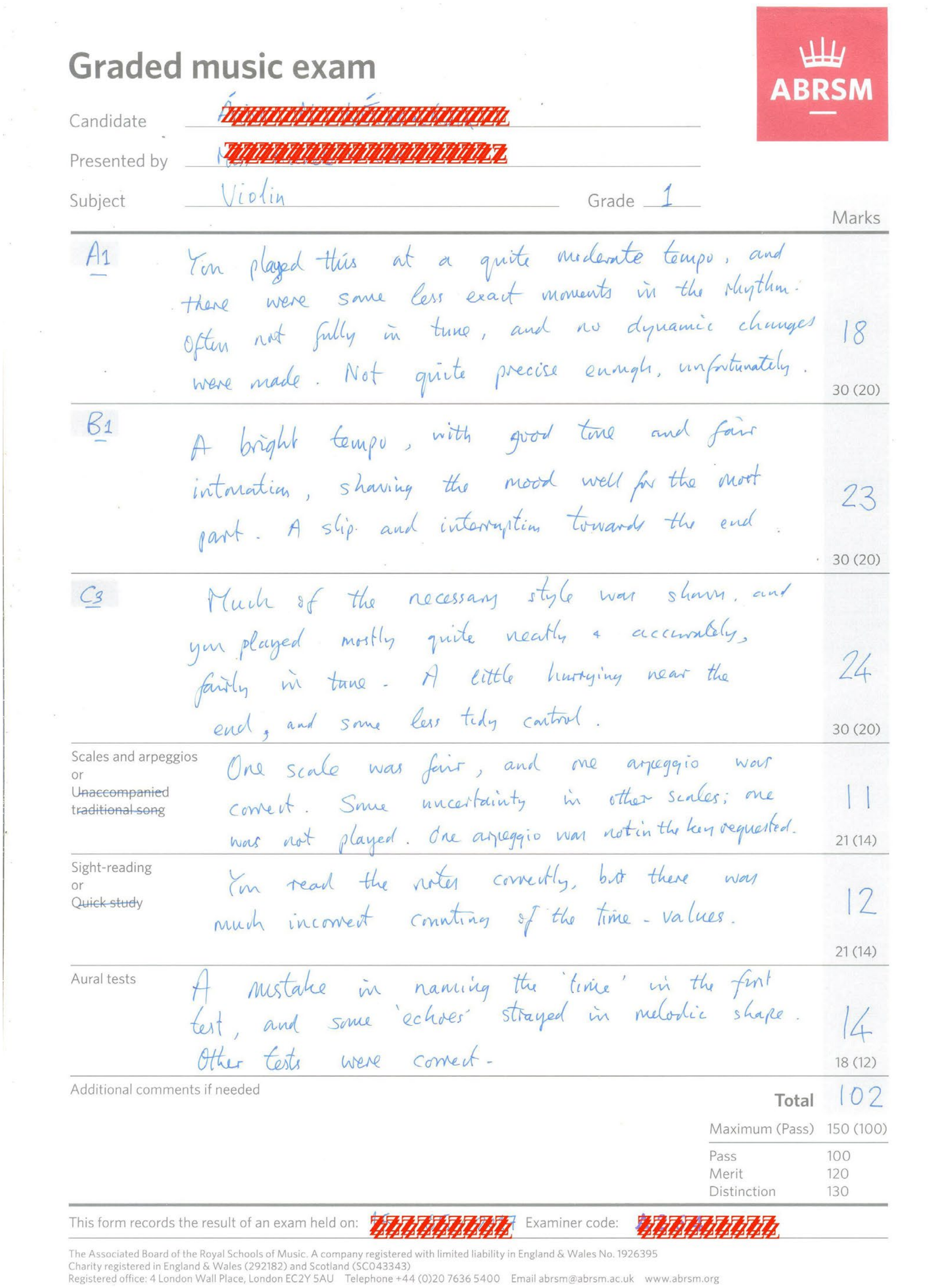
Figure 5. Analytic Rubric Wesolowski (2012)

In an attempt to address the shortcomings of rubrics, work has been undertaken into their further development and statistical validation. Ciorba and Smith (2009) for example, created a multidimensional assessment rubric, which was administered to 359 music students and used inter-judge reliability coefficients which demonstrated a moderate to high level of agreement among judges. This led them to conclude that "a multidimensional assessment rubric can effectively measure students' achievement in the area of solo music performance" (2009, 5). Similarly, Latimer et al. (2010) developed a multidimensional weighted performance assessment rubric, which demonstrated moderately high consistency, but which was within the range of previously researched music performance assessment tools. They found that "the rubric provided a better instrument for justifying ratings and more detailed descriptions of what constituted acceptable performances than previously researched non-rubric forms" (Latimer et al. 2010, 168). Aspiring to emulate and extend such research, the primary observation package (POP) developed for the study looks at candidate specific criteria - pitch perception, rhythm accuracy, music reading ability, and instrument specific criteria (strings) - intonation, tone production, bow trajectories and reading bow indications. This approach is more in keeping with a framework of educational Bildung, given that the feedback generated directly relates to learning milestones, which tend to be omitted in examiner generalisations and interpretations of a 'musical outcome'.

Building on these studies, in his later work, Wesolowski (2016) focused on rater precision in music performance. As was pointed out, the use of raters as a methodological tool to detect significant differences in performances and as a means to evaluate music performance achievement is a solidly defended practice in musical education, as the work of the ABRSM bears testament. In consequence, they examined rater precision through the analysis of a rating scale category structure across a set of raters and items within the context of large-group music performance assessment. Expert judges (N = 23) rated a set of four recordings by middle school, high school, collegiate, and professional jazz big bands. Their study found (p. 672) that "the data demonstrates that rating scale structures can vary by rater and each rater demonstrates unique tendencies of leniency/severity." Moreover, they state (p. 673) that "raters, regardless of 'expert status' do not share the same interpretation of rating scale structure," and "evidence that traditional estimates of rater consensus do not provide substantive meaning regarding the precision of true score performance estimates". Further work by Wesolowski (2019) analysed 1704 scores by 142 rates across nine high school solo and ensemble festivals. The results caused him to conclude that:

> the field of music education research can find great value toward the improvement of technical aspects of music performance assessments should scholars in the field decide to become more aware of … scoring automation research, and the applications of predictive regression and/or classification modelling such as machine learning drawn from the contexts of educational measurement research. (Wesolowski, 2019, 622)

Figure 6 is an example of a mark sheet for a participant in the study, indicating generic observations. These generalisations tell us little about measures which must be taken by this student to improve instrument specific technical obstacles, even though the student has entered a practical *violin* examination. Whilst ‘musical outcomes’ are the stated observation goal of the grade examinations model, a rigid adherence to instrument category on entry would suggest that instrument specific components should feature in some way.



Graded music exam ABRSM

Candidate [redacted]

Presented by [redacted]

Subject Violin Grade 1

		Marks
A1	You played this at a quite moderate tempo, and there were some less exact moments in the rhythm. Often not fully in tune, and no dynamic changes were made. Not quite precise enough, unfortunately.	18 — 30 (20)
B1	A bright tempo, with good tone and fair intonation, showing the mood well for the most part. A slip and interruption towards the end.	23 — 30 (20)
C3	Much of the necessary style was shown, and you played mostly quite neatly + accurately, fairly in tune. A little hurrying near the end, and some less tidy control.	24 — 30 (20)
Scales and arpeggios or ~~Unaccompanied traditional song~~	One scale was fair, and one arpeggio was correct. Some uncertainty in other scales; one was not played. One arpeggio was not in the key requested.	11 — 21 (14)
Sight-reading or ~~Quick study~~	You read the notes correctly, but there was much incorrect counting of the time-values.	12 — 21 (14)
Aural tests	A mistake in naming the ‘time’ in the first test, and some ‘echoes’ strayed in melodic shape. Other tests were correct.	14 — 18 (12)

Additional comments if needed

Total	102
Maximum (Pass)	150 (100)
Pass	100
Merit	120
Distinction	130

This form records the result of an exam held on: [redacted] Examiner code: [redacted]

The Associated Board of the Royal Schools of Music. A company registered with limited liability in England & Wales No. 1926395
Charity registered in England & Wales (292182) and Scotland (SC043343)
Registered office: 4 London Wall Place, London EC2Y 5AU Telephone +44 (0)20 7636 5400 Email abrsm@abrsm.ac.uk www.abrsm.org

Figure 6. ABRSM Exam Report.

Having considered the literature with respect to the ABRSM's grading system, and its use of rubrics, the following section assesses the literature in relation to the thesis's central theoretical backbone, and the development of Bildung, and its centrality to music education. The aspiration is to develop an historical analysis of Bildung, examining its variability, and provide some insight into the broad array of uses of the concept, and its renewed relevance in the contemporary discourse within music education.

2.9 Theoretical Framework: Bildung and Musical Education

The notion of Bildung as an educational theory and philosophy is often associated with the work of Wilhelm von Humboldt, although the concept pre-dates his analysis. Hence, the aim in this section of the thesis will be to examine the religious genesis of Bildung and its subsequent secularisation, and assess its relevance to music education. As shall be shown, using von Humboldt's notion of Bildung as an explanatory theoretical lens, with which to examine the raison d'être of musical education is a difficult, yet rewarding, task. In terms of difficulties, the first problem is that von Humboldt's exposition of this complex topic was exceedingly limited. Indeed, in the published volume of von Humboldt's letters, edited by Flitner and Giel (1995), the essay on the "Theorie Der Bildung Des Menschen" occupies only seven pages. Lüth (1998, 45) notes that "in November 1793 Humboldt complained that there was no more than an embryonic 'theory of human Bildung'" but his attempt to address this deficiency produced only a minor text, which Lüth (rightly) describes as "fragmentary". Hence von Humboldt's definition of the concept is neither thorough nor complete but in his 1793 fragment, dealing with Bildung, Humboldt (n.d., 284) made direct reference to the Germanic linguistic context of the term, stating that "but when in our language we mean bildung, we mean something both higher and more inward, namely the disposition of mind which, from the knowledge and the feeling of the total intellectual and moral endeavour, flow harmoniously into sensibility and character."

Secondly, as Varkøy (2015, 19) argues "the term Bildung has no direct counterpart in English," while Prange (2004, 502) "found as English equivalents to Bildung an impressive list of terms such as 'formation', 'growth', 'shape', 'training', 'education', ‘culture’; and ‘higher education’, ‘higher culture’, ‘refinement’, ‘good breeding’; and, correspondingly, in French: ‘culture’, ‘civilisation’, ‘formation’, ‘façonnement’, ‘discipline intellectuelle’. … The term covers a wide range of connotations and applications which are beyond definition.” Third, the genesis of Bildung drew on many factors, leading Horlacher (2004, 409) to draw attention to the “varied influences of a religious, literary theory, and aesthetics nature that give indications as to why the construct of Bildung has remained diffuse and excessive … which may be another reason for the difficulties that the German theory of Bildung continues to present.” Finally, as Kertz-Welzel (2016, 59) points outs, ethnolinguistically speaking, academic papers in German and English follow very different patterns. For the former, the author has “no need to explain everything, because the reader is expected to be an expert of some kind” while in the latter “the author not only describes facts and delivers information, but tries to persuade the reader of his or her own opinion.” Hence academic discussions in German and English on the topic of Bildung are likely to take very different approaches. Moreover, an English-speaking researcher into Bildung is further hampered, as the overwhelming majority of the literature on the topic is in German.

In terms of the genesis of Bildung, Alves (2019, 3) describes how the use of the term, “in the sense of the cultivation of the spirit, goes back to the 14th century Rhenish mysticism in which is designated the image of God that penetrates the core of the individual and thus shapes his soul,” indeed, the term was created by the German philosopher and theologian Meister Eckhart (1260-1328). Horlacher (2016, 8) describes how Eckhart translated biblical texts into German for the benefit of nuns who could not read Latin. The technical limitations of the German language were such that he was forced to invent new terms. Attempting to describe the process of spiritual transcendence into real humanity, achieved by contemplative self-devotion to considering oneself as made in the image of Christ, Eckhart used the term Bildung, and the Old High German word bildunga which, in its secular sense, had originally been used to signify the creative production of, for example, pottery. It is a linguistic paradox that Eckhart’s attempts to simplify and render explicable, from Latin into German, the process of religious transformation, gave rise to a term that is generally considered untranslatable from German.

The advent of the Enlightenment in the 17th century in Germany made it necessary to break free from the theological ideas of the uniqueness of the human being as God's creation, and the seeking of spiritual fulfilment and perfection, and turn towards empiricism. Chief among Enlightenment thinkers who assisted the subsequent process of the secularisation of Bildung were Locke, Rousseau and Shaftesbury. In the conclusion to his *Thoughts Concerning Education* (1693, 261) Locke conceives of children as "white paper, or wax, to be moulded and fashioned as one pleases," for whom the role of education is crucial, viz. ''I think I may say, that of all the Men we meet with, Nine Parts of Ten are what they are, Good or Evil, useful or not, by their Education. 'Tis that which makes the great Difference in Mankind" (Locke, 1693, 2).

However, as Crittenden (1981) makes clear, Locke's idea of human perfectibility transcends education, in the sense of the acquisition of useful or utilitarian knowledge; rather, his emphasis is on the acquisition of virtue, the formation of moral character. Thus, in section 134 Locke (1693) lists virtue, wisdom, breeding and learning as the cardinal aims of education and hence proclaims: "I place Vertue as the first and most necessary of those Endowments that belong to a Man or Gentleman" (section 135). "Learning", he explains, "must be had, but in the second place, as subservient only to greater Qualities" (section 147). Similarly, in section 70 he states: "Tis Vertue, then, direct Vertue, which is the hard and valuable part to be aimed at in Education."

Rousseau (2002, 96) also acknowledged the possibilities of human self-improvement towards a state of perfectibilité, by noting that "there is another very specific quality that distinguishes them, [i.e. humanity] … this is the faculty of improvement; a faculty which, as circumstances offer, successively unfolds all the other faculties, and resides … in the individuals that compose it." Additionally, Horlacher (2016, 23) notes that Rousseau's educational novel *Emile* (1762) "combines a concept of human perfectibility with a belief in the practicability and predictability of education, all of which ought to provide a sound basis for a society in which individual and social development are correlated."

Publications by Anthony Ashley Cooper, 3rd Earl of Shaftesbury (1671-1713) provided a further English contribution to the philosophical foundations of secular Bildung. John Locke was Cooper's tutor, and Cooper's education by Locke was in close accord with Locke's beliefs as espoused in his treatise *Some Thoughts Concerning Education.* In Cooper's 1710 essay, *Soliloquy: Or Advice to an Author*, he argues that the crux of dispensing sound advice is for the author to acquire self-knowledge, such that the advice he dispenses is accurate, unbiased, and not subjective. Hence Cooper (1710, 16) advises the reader that "unless the Party has been used to play the Critic thoroughly upon himself, he will hardly be found proof against the criticism of others. His thoughts can never appear very correct unless they have been well informed and disciplined before they are bought into the field. Tis the hardest thing in the world to be a good Thinker, without being a strong self-examiner, and been a thoroughly paced dialogist, in this solitary way."

In the view of Rowson (2019, 5) "Cooper was the first to emphasise the importance of 'inner Bildung', our inner formation, not merely for its own sake, but because the nature and quality of our inner formation (and realisation) is reflected in 'outer Bildung' in the systems and structures of society, and their nature and purpose." Indeed, Horlacher (2012, 138) notes how Cooper "describes this self-examination, and purification process at length describing it using the words 'to form' and 'formation'. In the German translation of 1738 [by the theologian and philologist, Georg Venzky], these words were translated as bilden and Bildung: they were the subject of intense discussion."

Assessing the contemporary context for Bildung, Alves (2019, 2) correctly, but unhelpfully, surmises that "the concept of Bildung is one of the fundamental concepts of modernity and the most ambiguous concept of German pedagogy, providing a range of uses and interpretations." However, Zelić provides an accurate and succinct description of its development and promise:

> Bildung "does not prescribe the acquisition and transmission of a given body of empirical knowledge about the historical facts of nature and culture, but aims rather at self-development and perfection of everyone's individual personality. Bildung, accordingly, does not aim primarily to educate individuals to satisfy their material needs and wants. It is rather an infinite task and permanent assignment of self-reflexive and critical understanding and endowment of meaning to the self, others, and the world. It requires proficiency in many different languages and cultures and their integration into the totality of one's individual personality, which will enable individuals to participate in the exchange of

> innovative ideas and the critical dialog in and between cultures, to discuss the most important questions of humanity in the past, present, and future in order to improve the power of judgement and to act accordingly". (Zelić, 2018, 664)

Despite Bildung's amorphous qualities, various authors have argued that it has direct contemporary resonance with musical education. Heimonen (2014, 190) for example, considers Bildung to be "part of a holistic view of music education" and argues that it may refer to "an individual's spiritual maturity that she or he has received from a broad education, especially including the arts and literature, or without formal knowledge from practical life experiences" and that it "also has a collective meaning that is connected to the Bildung of a nation or even of larger communities, for instance, 'Western' or 'Eastern,' that comes close to the concepts of 'culture' and 'civilization.'" She further argues (p. 195) "Education that is based on the individual needs of every student, an ethos that 'music is for all' and that 'everyone has a right to music,' is closely connected to a view of Bildung; Bildung holds the view of a broadly-educated human being, not only literally but also practically". Similarly, Kertz-Welzel (2017, 109) maintains that "Bildung is the core idea in music education and music education policy … First, Bildung in music is part of Bildung in general, supporting the development of a cultivated person; this concerns non-musical goals such as fostering intelligence or creativity through music. Second, there is a specific Bildung in music in terms of gaining musical knowledge and skills."

In respect to the analysis of Bildung, Frede Nielsen was one of the most prominent European philosophers in music education; however, although the author of many papers, his influence has been limited outside of his native country of Denmark, owing to his disinclination to write in English. Nielsen (2007, 269) describes the nature and goal of Bildung as follows: "The goal is man's self-determination and autonomy based on reason, a life in freedom and mutual respect between fellow human beings." Although his analyses of Bildung are thorough and comprehensive in scope, Nielsen did not see his work as prescriptive, indeed he states (2005, 7): "it is not my intention to tell what music teaching and learning ought to be like, but to offer others improved possibilities of making conscious and reasoned subject-didactic choices." In consequence, much of his work is descriptive and analytical, rather than normative which, however, lessens the utility of his work to music educators who wish to apply

Bildung to their teaching. For example, in his analysis Nielsen (2007, 270) distinguishes between Material Bildung theories, in which "the teaching/learning matter, including the subject-related forms of activity, is central and in itself a criterion for the selection of educational content" and Formative Bildung theories, in which "the person going through and becoming formed by the process of Bildung ('formation') is in focus. The educational content is the means to realize this formation." Hence, Material Bildung theories identify which aspects of our multifaceted reality are so valuable that a student should learn or experience them, while Formative Bildung theories ask which behaviours, competencies, and methods will be important, in the present and future. Further, Nielsen conceptualises Bildung thus:

> The idea is that human beings do not have personal individuality when they are born. This is gradually acquired in a process of Bildung that leads to personal freedom. Bildung must be complete in the sense that all of a person's powers ("Kräfte"), not just single skills, should be cultivated. The attainment of personal individuality means that the powers are developed as an integrated whole. The content of Bildung is based on this aim. It should be universal in the sense that it represents the spiritual structures and values that are necessary for the individual's complete development. It is thus "general" in the double sense that it has the status and character of something that is both of a general sort and of significance to everyone. (Nielsen, 2005, 269)

All of the main literature on musical Bildung by Nielsen (and others) relates to students' Bildung. However, Heimonen also conceives of Bildung as an expert culture of music teachers as well. According to Heimonen (2014, 196) music teachers embracing Bildung adopt "a view of education that aims at teaching to respect others' cultures, including teaching pupils to understand that music has different kinds of meanings for various individuals and that all these meanings and different kinds of musics are equally important and valuable." Furthermore, such teachers follow "a dialogical way of teaching and a democratic relationship between the educator and the educatee," in which "the aims of education are negotiated collectively between the pupils and the teacher, and … democracy in music education refers to a respectful and tolerant atmosphere, in which pupils are educated towards becoming autonomous, critical, and active members of society." Hence, in the sense of self-fulfilment of personal, social, cultural and democratic potential, music teaching aids the development of Bildung in teacher and student alike. Bildung requires self- reflection as a way of professional improvement and personal growth. Currently the ABRSM examination does not allow for this, as it is based on a single performance, rather than an on-going process. However, as the thesis will demonstrate, new technologies enable the capture of individual data with respect to instrument playing, which allows students to see themselves, for themselves, and reflect on how they are playing, and thereby enable music teachers to suggest adjustments to improve the technical proficiency of students. In this way, the underlying rationale shifts from being an assessment of a performance, towards a process of continuous improvement and the creation of Bildung for both pupils and teachers.

2.10 Summary

The literature review started with an examination of string performance assessment, and the initial work of pioneers in the field, namely Seashore (1915) Wing (1962), Gordon (1965) and Bentley (1966), whose work together constitutes the foundations of contemporary music assessment. More recent work has shown that assessment instruments to measure musical attainment vary in terms of their reliability and validity; moreover, even when these are not in question, rater evaluations can be highly inconsistent. Further work on refining evaluator mechanisms was shown to have given rise to two main views on assessment. The first conceives of measurement being anathema to musical performance; by contrast, the second sees the testing and standardisation of music performance as being both normal and desirable. The literature review demonstrates that no previous study has examined all of the assessment criteria in the current study.

Moving on to consider the development of frequency change detection and intonation, the literature review addressed the work of Helmholtz, but which was refined and augmented by Ellis. In terms of the detection of frequency change detection and intonation, the literature reveals that young children are articulating tonal awareness, can discriminate quarter tone intervals, and identify diatonic structure. Moreover, research reveals that timbre was found to affect the ability to detect discrepancies in pitch frequency, and is also related to the fingers used to play notes with computer models now being able to anticipate the most suitable ones to be used. This element of the literature review reinforces the rationale for the first two of the ten variables, namely, frequency change detection and intonation. Similarly, the previous work on measuring bowing posture and tone production has been shown to have been radically improved through the innovations by Ng et al. (2007), Schoonderwaldt & Demoucron (2009) and van der Linden et al. (2011a). These findings show that growing use of technology in musical teaching has enabled greater attention to be paid to teaching strategies which can improve posture, thereby enhancing musicianship.

With respect to assessing musical literacy acquisition, the review examined the literature in relation to expected parameters and current norms which inform the design in the thesis in relation to the literacy variables of note pitch, metric duration and bow direction indications. Surveying the literature revealed a significant gap in the research literature, in relation to reading and the assessment of bow indications. Moreover, various contemporary authors in the field such as Tomlinson (2015) and Burton (2015) have called for a revaluation of the role of literacy in music instruction, making this current study both apposite and timely.

The review then examined the use of emerging technologies in music education in more depth, starting with the prescient belief of Seashore (1938, 30) "musical aesthetics will be built on the basis of scientific measurement", and examining how successively more sophisticated technologies, and their related music information retrieval softwares are bridging the gap between music education and computer science. The review demonstrated that there has been a steady growth in the number and size of projects looking to bring the computer into the musical education classroom, many being international in scope and supported with EU funding. Arising from the review, it was decided to use Kinovea and Melodine software in the study – the attributes of these programmes are considered in more detail below.

The literature review revealed general dearth of published research examining the validity and reliability of the ABSRM assessment and grading system. Moreover, it was noted that, in the absence of such studies, the Board's unwillingness to allow findings from the only detailed evaluative statistical study of the Board's grading and assessment system to be cited, more than 20 years after it was undertaken, tends to undermine rather than reinforce, the credibility of the Board's system of grading. Following from that, the literature investigated the features of marking rubrics, (as are used by the ABSRM), using Wesolowski's models of holistic and analytic rubrics, and assessing their lack of objectivity and their limited abilities at providing practical formative feedback. Work by Ciorba and Smith (2009) and Latimer et al. (2010), on multidimensional rubrics, shows how more sophisticated forms can provide better grading consistency, which is the aspiration of the current study. Moreover, such a multidimensional approach, through the technological capture and study of features of string playing, is designed to build attributes of personal confidence, technical competence and musical mastery, through a Bildung approach to music education.

The final section of the review examines the history, development and centrality of the German philosophical and pedagogical concept of Bildung to contemporary discourses on music education. Unfortunately, because the virtually all of the voluminous work on Bildung is only available in German, this section of the review provides only a partial examination of the discussion surrounding Bildung. Nevertheless, despite the limitation imposed by just examining the discussions of Bildung in English, it is evident that, as both a concept and a pedagogic aspiration, Bildung has a particular resonance for music education. It is clear that achieving Bildung is a personal journey for music students, seeking to improve their technical expertise and, thereby, enhance their public performance playing, such that it pleases both pupils and their audiences. The methodology adopted in the thesis offers the opportunity for pupils to engage in self-examination of the physical attributes associated with their string playing and, tutored by their music teachers through a process of self-reflection, to achieve mastery of their instruments, and realise their full potential as musicians. Similarly, by enhancing the resources available to music education, the use of new data capture technologies as proposed in the thesis, will broaden, expand and enrich, the roles of music teachers, thereby enabling them to achieve Bildung as educators.

In sum, the literature review addresses the limitations of the current method of musical assessment and grading, as exemplified by the ABSRM approach, and demonstrates clearly how the proposed research will utilise previously unmeasured observation variables including frequency change detection, intonation accuracy, bow trajectory, posture, tone production and music literacy acquisition. Emerging technologies and innovative ways of incorporating them within these observation criteria have been explored. The ABRSM musical outcomes assessment model (and the underlying rubric system) has been discussed and the central role of Bildung in relation to music education explained.

In the next chapter, changes to the pilot study are explained and the observation schedule is detailed. The operationalisation of the study is introduced detailing software choices, data formats and observation criteria. The assembly of ten observation variables is then explained, giving for each variable, theoretical underpinnings, related work, technical means of recording the data, calculations involved, observation protocols, experiences of participants, and examples of how feedback was returned.

3 Chapter Three Methodology

3.1 Introduction

The aim of the research is to make a positivist, data-driven comparison between participant ratings in 10 objective variables, relating to instrumental performance, and standard subjective grade results. 37 violinists participated fully in the research, with half coming from private tuition backgrounds and the others originating from classroom instruction settings. All participants underwent the same observation procedures and the same type of grade examination process. The comparison shows a link between objective and subjective assessment.

While surveying the literature, to better understand constructs that already exist in rating scales for string players listed by Zdzinski and Barnes (2002), it is born in mind that the aims of the thesis do not include replication of these constructs. However, theoretical support is given to constructs brought about by technological advances which do not feature in the existing literature. In addition, technology has opened up new possibilities and ways of gathering data to determine patterns in performance with a high degree of accuracy which was unavailable previously. Determining, validating, and correlating these new constructs creates new knowledge in a field that is generally surrounded by ambiguity and uncertainty on the one hand, and technical complexity on the other. The parsimony of variables chosen, and the logic behind their association with one another, will be expanded on here.

The degree to which test structures vary affects the accuracy of responses. A feature built into the design of the variables is one of synchronous interaction. This happens where the responses are compiled and documented together with other relevant material. For instance, transcriptions of notes played in the intonation test are presented graphically, with frequencies of the notes played along with video material. This feature in the research design makes verification and repeatability more exacting and transparent. Similarly, bow distances are plotted and superimposed on the visual image. Quantification of elusive bow trajectory and its distance travelled, can therefore be determined and measured successfully, thereby bolstering construct and content validity.

3.2 Findings of the Pilot

The pilot looked at intonation, as measured by Melodine software, and bow distance travelled, as measured by Kinovea, to understand the relationship between these two aspects of violin playing. It was postulated that when intonation was poor, a common response is to curtail the amount of bow being used to minimise the negative effect. The indicators gathered from the initial study confirmed this, indicating that bow distance (producing a strong tone quality) is compromised when participants play out of tune. Scores from intonation tests, and scores from bow distance travelled observations, were made into percentages and analysed with the Pearson correlation statistic.

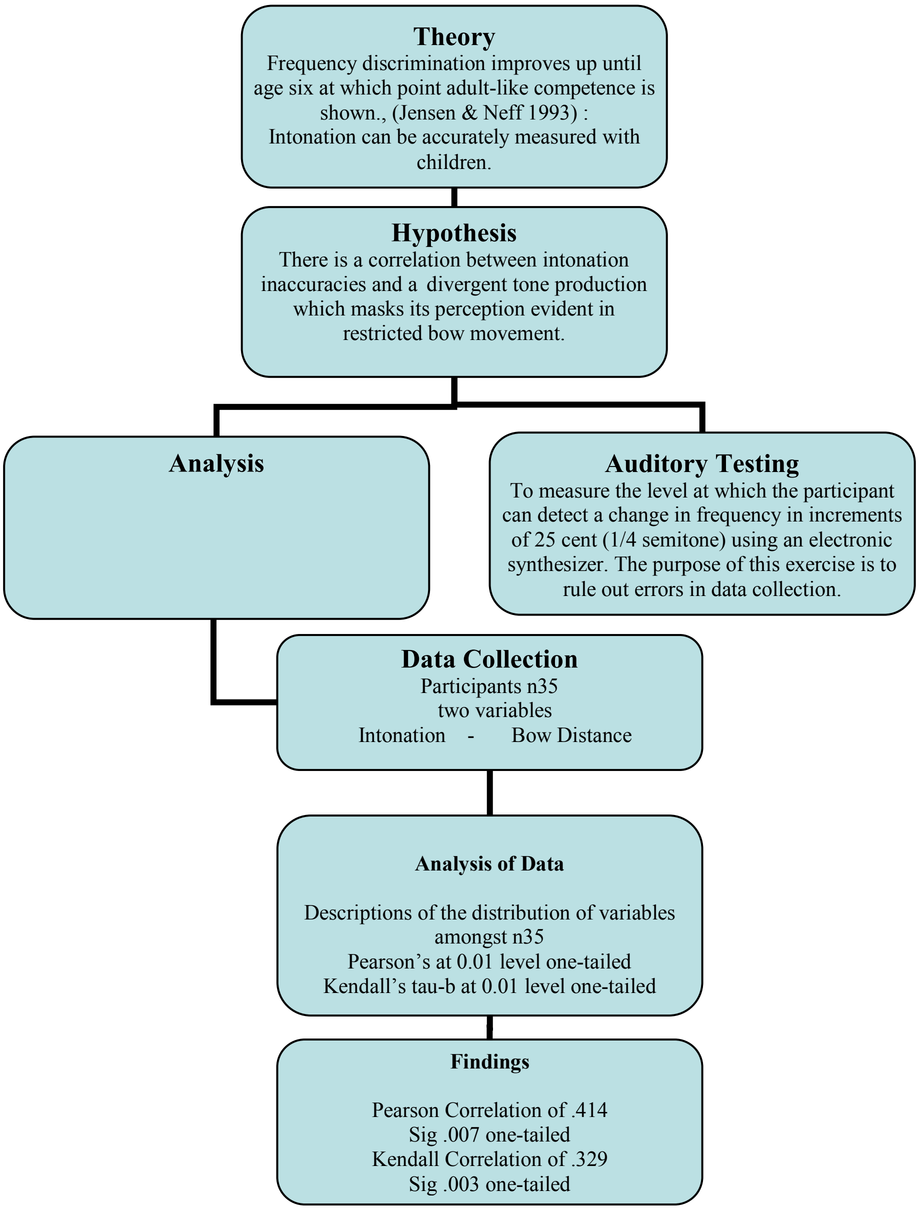


Figure 7. Pilot Study

The pilot study diagram (in figure 7) shows the method by which a hypothesis was tested to ascertain if there was a relationship between intonation accuracy and the amount of bow that a student used. The assumption underlying this being that when a student plays notes which are out of tune, the amount of bow being used is reduced, to counter the effects of out of tune sounds on the performance. This tendency became apparent over many years of teaching and the phenomena was of particular interest because an improvement in intonation did not automatically lead to greater bow use. In other words, intonation issues seemed to cause a restricted bow movement which can remain a problem after the intonation issues are resolved.

The theory suggested that students should be capable of responding to intonation issues directly. The hypothesis postulated a link between the two occurrences. Auditory testing was undertaken to ensure that all students were hearing the sounds correctly. Thirty-five students were observed with the protocols described in the following section, in relation to intonation and bow distance. A Pearson and Kendal test was run, and a moderate to high significant correlation was found to exist between both variables.

	Intonation Percentage
Distance Percentage	r =.414
Significance (1 tailed)	.007 (Significant at 5%)
Sample Size	35

Figure 8. Pearson Correlation in pilot

The outcome of the pilot signalled a moderate to high correlation between the two variables. The pilot was then expanded to take account of a broader range of instrument specific variables – pitch detection, intonation, rhythm, instrument angle, trajectory, distance, guidance, reading and grade result. The results are explored in the following pages, which address the research question, which postulated a link between these observed variables and grade outcome.

The pilot study helped to streamline the data gathering process. Given that many observations were required to gather data for each variable construct, observable instrument specific singularities were given priority. The design adopted also made possible concurrent validity, as per Muijs (2004), enabling crossover and comparisons of sub-field ratings, meaning that individual variables could be correlated with one another, (as is shown in Figure 25 on page 133). Five observation sessions recorded data relating to the 10 variable constructs to test the hypothesis and address the first research question which was: Does an instrument specific approach to performance assessment correlate with musical outcomes in practical ABRSM grade examinations? The schedule developed from the pilot study is described in Figure 9 on page 76.

3.3 The Operationalisation of the study

The study utilised four different computer programs. First, the audio analysis was achieved with Melodine software, which is described in detail by Hoenig et al. (2015). Generally used for pitch correction in the recording industry, this innovative application differs from other pitch recognition software, in that it has the ability to access polyphonic note frequencies at a post-production stage, referred to as direct note access, or DNA. The software was incorporated in the study because of its accuracy in locating and determining exact pitch frequency discrepancies. The software has considerable academic credibility, having been cited in studies relating to high-level audio signal transformation by Yeh et al. (2010), pitch trajectories in speech processing by Järvinen-Pasley et al. (2008) and vocal intonation by Hutchens and Peretz (2012). The many features of the software include comprehensive visual representation of the sound, with discrimination of pitch with 1 cent accuracy (1/100 of an equally tempered semitone).

Secondly, motion capture measurements were obtained using Kinovea open source software. As with Melodine software, this programme has been used widely for academic research. So, for example, amongst its many applications, this software has been used for lower limb analysis in Guzmán-Valdivia et al. (2013), sports rehabilitation in Bačić (2015), and martial arts training in Branco et al. (2016). The relatively new application of this technology to analysis reflects a growing need for measuring verifiable data in relation to movement. A comprehensive study of its development is described by Moeslund et al. (2006). These two software technologies made observations possible without wires or restrictions being placed in or around the participant's immediate performance space, making data collection possible without participants being restricted in any way, or the need for observations to take place in a laboratory environment.

Thirdly, Sibelius is a music notation program, which produces printed scores but can also play the music back using sampled or synthesised sounds. Finally, Garageband is a digital audio workstation that enables users to create multiple tracks with pre-made MIDI keyboards, pre-made loops, an array of various instrumental effects, and voice recordings

3.4 Observation schedule

Five observational sessions took place over five consecutive mornings between January and February 2016. Each session took approximately three hours, giving five minutes observation time for each of 37 participants. The total observation time took 15 hours. A further 30 hours was spent synchronising the motion capture data to make posture determinations and distance and direction measurements. Transcriptions took a further 10 hours to prepare data for analysis. Figure 9 details the process of the data collection schedule.

In order to ensure transparency and clarity, the methodology chapter describes and explains the observation process through which the measurements in the study were made, by giving examples of participants who took part, and detailing implementation procedures and protocols followed at each stage of the observation process. Attention is given to the way in which the variables were developed, the processes involved in measuring them, and the level of precision achieved. The following table (in Figure 10 below) indicates the observation protocol for each variable, the format adopted and the software used to operationalise the study.

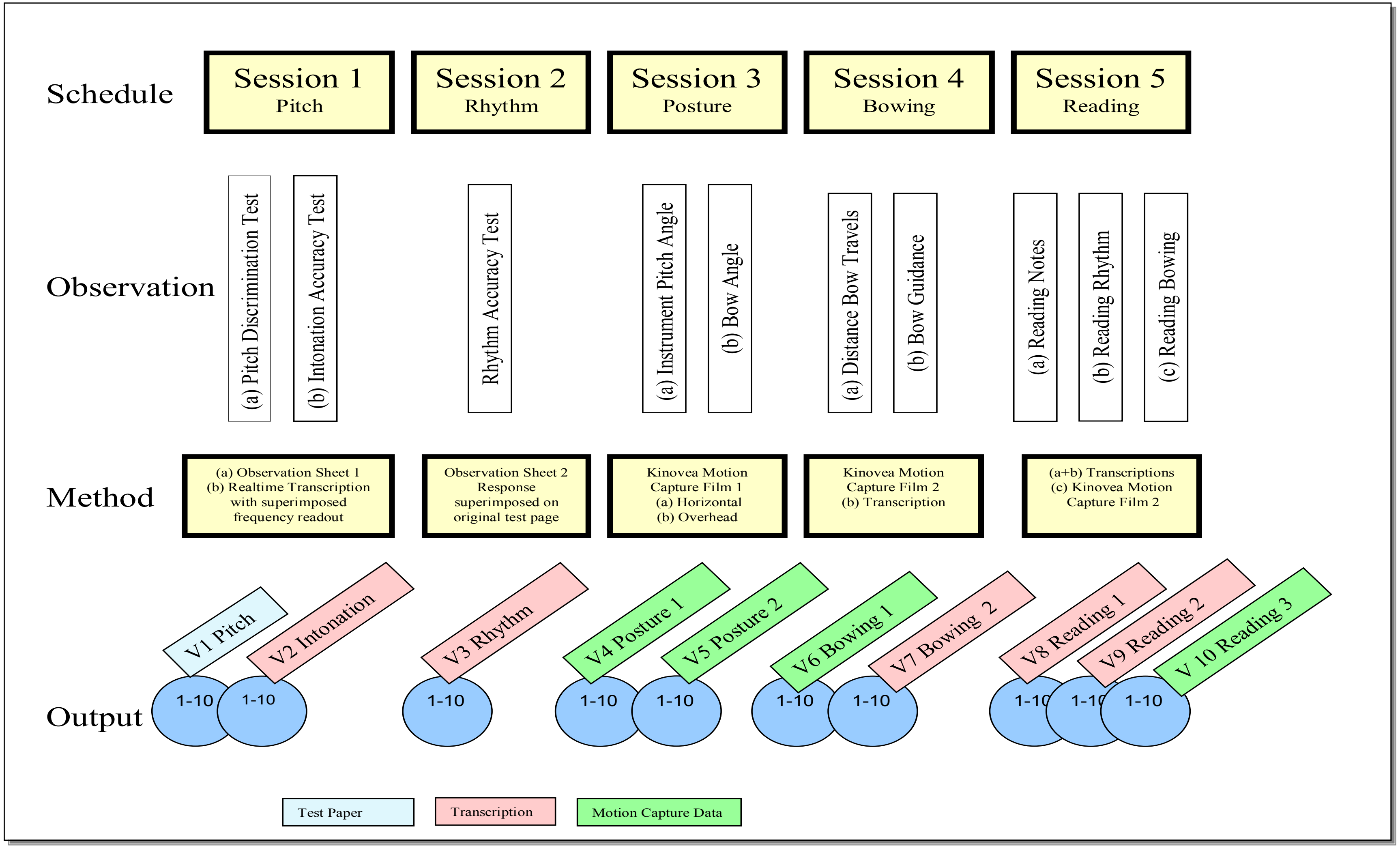


Figure 9. Observational Schedule

Variable	Protocol	Format	Software
1- Pitch	Questionnaire	Written	Sibelius
2- Intonation	Scale of D Major	Audio	Melodine
3- Rhythm	Response Test	Audio	Garageband
4- Instrument Angle	Own Choice	Video	Kinovea
5- Trajectory	Own Choice	Video	Kinovea
6- Distance	Scale of D Major	Video	Kinovea
7- Guidance	Sight Reading Extract	Audio	Melodine
8- Reading Notes	Sight Reading Extract	Audio	Melodine
9-Reading Durations	Sight Reading Extract	Audio	Melodine
10- Reading Bowing	Sight-reading Extract	Video	Kinovea

Figure 10. Processes of Data Collection

3.5 Observation structure

Each observation section begins with a short description of the background to the variable concerned, its value, in terms of the theoretical contribution it makes to overall performance ability, what is known in the literature about similar observations, the technical means of recording the data, the protocols followed during the observations, the calculations made when processing the raw data and each section concludes with a summary of how the information gathered for each variable was returned back to the site of learning, in accordance with educational Bildung theory mentioned earlier. In relation to music assessment and standards, Fautley and Colwell (2018) state:

> 'One assigns value, one describes the meaning of the data and observations, one synthesizes experiences, and the resulting judgments indicate the merit, worth and significance of the educational venture' (Fautley and Colwell, 2018, 258)

This approach is at the heart of the observational process at hand, and for each of the observation variables, the discussion has been structured under the following headings.

- **Theoretical underpinnings which justify the variable**
- **How the variable sits with related work**
- **The technological means of recording data**

- **Calculations required to normalise the data**
- **Observation protocols and experiences of participants**
- **Examples of how results were returned to the learner**

3.6 Variable 1 – Pitch Discrimination

3.6.1 Theoretical Underpinnings

Gingras et al. (2015) estimated that about 3% of the population have difficulty detecting notes that are out of tune. The ability to play in tune is affected by the ability to determine if something external is in or out of tune. This ability is related to culture, context, faculty and pedagogy. Tan et al. (2014, 15) found "the intriguing possibility of an overlap in the neurobiological basis of music functions and social behaviour". The extent to which social behaviour speaks back to pitch detection depends on the value and cultural norms placed on its development. Western music culture is, to a greater or lesser extent, reliant today on adherence to equal temperament. The various genres, which share a place in music education, apply pitch discrimination in learning unevenly; however, Western art music culture places more emphasis on this aspect of pitch discrimination than others, as it underpins the cultivation of intonation awareness.

Equal temperament, on which Western music is based, is actually an 'out of tune' phenomenon as perfectly in tune intervals of a fifth are actually tempered 'flat'. This is done to compensate for the unsettling reality that when twelve intervals of a fifth are stacked together – for instance, above the given note A 220 Hertz – the arrival on the same letter name sound, twelve (fifths) steps later (descending an octave where required) – that is, to arrive on A 440Hertz – this new note will then be unusable, as it will be too sharp.

The margins in which equal temperament differs from 'just temperament' (as it is called when intervals are perfectly in tune) can be quantified by the division of the scale in 12 equal parts, and then further dividing semitones into 1/100 parts, or cents. This idea was first introduced by Alexander J. Ellis who, despite not 'trusting' his own ear, put forward the cent system for measuring pitch discrimination abilities and sensitivities in others in Ellis (1876, 31). He states that "A good ear is one which, within the 6th, 7th, 8th, and 9th Octaves appreciates, both in distance and direction, an interval of one or two cents in Unisons, Octaves, and Fifths, and ten to fifteen cents in other intervals" (p. 24). It is an unrealistic expectation to anticipate this level of discrimination from novice players, in my view.

In order to operationalise the study for the observation of novices, therefore, it is postulated that a generous margin of 29 cent outside of a fundamental frequency, sharp or flat, (that is about 1/3 of a semitone above or below the fundamental frequency) should be easily perceptible by novice participants. Beyond this bandwidth of intonation tolerance, a tonal framework from which music ideas may be constructed and shared in the Western tradition would become unworkable. That is, for this study, more than 29 cent above or below the note is considered to be perceptibly out of tune, in the pitch discrimination test. Sloboda (1985) used the term "generous minimum" as a better fit, in making useful determinations about novice string players' abilities to regulate pitch when reviewing earlier test models. Three of the 10 questions in variable 1, however, had adjustments smaller than this amount made to the semitones (within Ellis's cent framework). This was undertaken to take account of exceptional ability, and it was anticipated that this less generous discrepancy minimum, set at 10 cent, would not be perceived by most of the participants. However, a small percentage of participants, (4 out of 37) managed to detect two of those three, receiving scores of 90%. These students are the exception. The test is, first and foremost, designed to ascertain if students can detect if something is fundamentally out of tune, more so than if they can detect if something is slightly out of tune. The latter aspect being concerned with test sensitivity and calibration, rather than data in that regard.

3.6.2 Related Studies

Cooper (1994) found when using a modified version of the Pitch Discrimination Measure (MSPDM) created by Sergeant (1973) that students were better able to detect differences in pitch than the direction of the difference. Norris (2013) conducted a tonal awareness study focussing on the perception of tonal dissonance. This study, looked more broadly at dissonance present in 26 items referring to them as 'clinkers'. Students were asked to determine if such notes had occurred in given extracts. The research suggested that experience, rather than development, is key to how children perceive musical information. Fancourt et al. (2013) used the term 'odd one out' when directing a similar study, which also looked at pitch discrimination and pitch-direction. His findings suggest that the two problems, that of pitch change detection and the understanding of the direction of that change, are separate cognitive tasks to be addressed independently.

Sun et al. (2011) found students may not be able to readily associate sounds they are hearing with a particular tonality, as in pitch class distribution theory. This may be because pitch discrimination can be influenced by the timbre of instruments as discussed by Geringer et al. (2015) who discovered that participants found it difficult to detect discrepancies in voice, compared to discrepancies in trumpet, even though identical music material had been presented in the test. The Hungarian music teacher Kodály favoured an acoustically 'pure' environment for teaching intonation through singing, because of the negative influence of equal temperament on pitch perception, due to the tempering on keyboards of what would otherwise be pure intervals according to Jaccard (2014). Although this is a relevant consideration, regarding orientation to a pitch discrimination test, it was considered outside the scope of what was being measured in the variable 1 test. This is because the test is looking exclusively at students' ability to detect if something is clearly 'out' of tune.

As would be expected, Hutka et al. (2015) found pitch discrimination is better developed in musicians than in non-musicians, yet maturation in frequency determinations is not dependent on temporal cues according to Buss et al. (2014). The ability to detect pitch change may follow a separate developmental trajectory from one needed to understand the direction of that change as found in Fancourt et al. (2013). While conflicts can invariably arise between subjective and fundamental frequency-based pitch assessments, as described in Vurma et al., (2011), Fautley and Colwell (2018, 271) provide a way forward by stating that it is helpful when assessment criteria 'relate to a singularity.' Focus on singularities; therefore, have influenced construction and implementation of each observation variable.

3.6.3 Recording of Data

The current study incorporates an 'adjusted' piano sound, to overcome the problem of timbre which would have adversely affected the reliability of the pitch discrimination test. The samples were recorded from a Yamaha Portable Grand

DGX-630 set on a ‘timbre proclivity neutral’ basic piano sound. This device had the ability to adjust selected notes temperaments in cents.

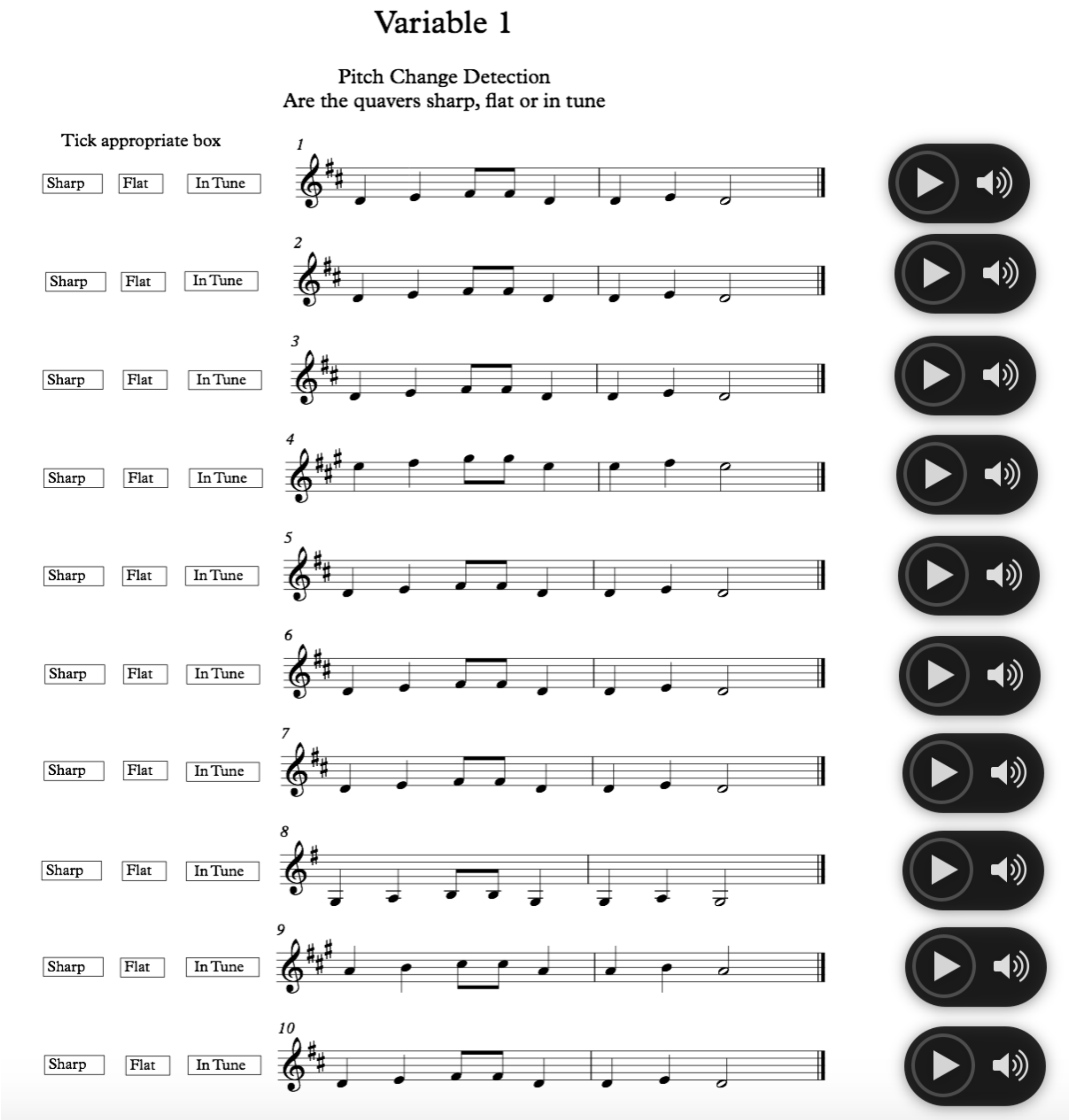


Figure 11. Pitch Discrimination Test

Essentially, the test consists of a 10 - part questionnaire produced with Sibelius software that is presented to the participants on arrival, who then listen to 10 short recurring extracts and comment on each one. The sound files were attached to a single page and could be played as required on a MacBook Pro computer and heard through headphones. As extracts 1 - 10 are played consecutively, the participant marks the score sheet for each extract, outlining if it is sharp, flat or in tune. The participant dictates the pace of the test and is granted multiple attempts at listening if required (See Figure 11. above). As was explained to the participants beforehand and on arrival, what they are listening for, is the two quavers (the 3rd and 4th notes in each extract) and they are then making determinations about these two notes only, as all the other notes are left in tune. This helps the students to focus their attention exactly on the place in the extract where the pitch adjustment may exist. Both quaver notes in each extract are altered in the same way.

Tonality in the test is simplified, in so far as the extract in each case begins on the tonic (key note). This is helpful as the expectation of the familiar motif in a major key enables the listener to be more objective during their listening, at the place where the adjustments have been made. They are expecting a regular major scale type motif to open each extract, and can anticipate how the extract will proceed, as it is the same repeated motif in each instance which is relatively timbre neutral. Neutral in this regard relates to the fact that some instruments are easier to notice being out of tune than others. This problem has been overcome with the use of a prepared digital piano, which maintains correct intonation (equal temperament) throughout, *except* for the adjusted quavers in the extract. Clear margins in 7 of the 10 extracts determined at 29 cent above or below the fundamental or, perfectly in tune, i.e. Sloboda's 'generous minimum' mentioned earlier, made it possible to say, with some degree of certainty, if the participant could respond to the elemental question at the heart of the observation, which is, can you tell if something is out of tune? And, if so, can you say in which direction it is out of tune, i.e. is it sharp or flat? Three answers were more fine-grained, being set at 10, 20 and 20 cent deviations.

No machine measurements were involved in gathering data for this variable which was assessed on the basis of what the participant had written on the page. A fully correct score, as written by a student, would have read - in tune, sharp, flat, flat, in tune, flat, sharp, sharp, in tune and sharp.

3.6.4 Calculations Involved

There are no calculations involved as 10 possible correct answers (raw data maximum of 10) accrue a score of 100%

3.6.5 Observational Protocols

The school music room was the site for the classroom observations and the private students, as teaching took place at the same location for both cohorts, the other being after school hours. Students were enthusiastic about participating in the test, as they were genuinely curious themselves to know if they could tell the difference, or what that differences might be. Classroom students arrived in groups of two, and as one participated in the test, the other waited for their turn. In some instances, the student who went first asked to re-take some of the extracts, which they were unsure of first time around, and this was allowed. Private students arrived unaccompanied. As it is not a sight reading test, students were encouraged to listen a second time if they felt that this would help them to make their determinations. However, they were not allowed to confer or compare or discuss their interpretations, so as to maintain the autonomy and validity of the sample. No feedback was returned about responses during the test.

3.6.6 Feedback to Learner

On the occasion that the variable 1 observations took place, it was not discussed with the students how they scored in the test. This information was made available to the students at a later stage, when the results were being discussed in a separate learning module. Some students chose not to be given their scores, with others being anxious to know exactly how they did. The results, however, form an important learning step in ascertaining if the student is well positioned or not, to address the intonation qualities of sounds made by themselves, that is, making adjustments to discrepancies which invariably exist when young novice string players start to develop listening skills associated with their own playing. Of the 37 participants who took part, four of them got a score of 90% and two participants obtained the lowest score of 30%. From this, it could be extrapolated how these students would do in the intonation test, as being able to detect if something is sharp, flat or in tune is a prerequisite for correcting one's own intonation. The observations above generated 37 score sheets with varying degrees of perception displayed. This data would constitute factual measurements, incorporated in the comparison study with grade outcomes in relation to the pitch perception singularity, as described in variable 1. This information was also made available to the classroom teacher, who could take meaningful steps to help students who were struggling with this particular aspect of musical comprehension.

3.7 Variable 2 - Intonation

3.7.1 Theoretical underpinnings

The perception of two notes being out of tune with one another, happens as a result of the perception of waves of similar frequencies beating together. Rapid beating equates to something being very out of tune, whilst slower rates are perceived as acceptably in tune. This has implications in music performance, and comes under the broad heading of intonation when playing a stringed instrument. In relation to novice string learning, it involves establishing in the learner an ability to create semitones with the fingers of the left hand, which are a small enough distance apart to be called semitones and tones, which are far enough apart to be perceived as tones. There is a tendency for novice players to make different types of intervals the same size.

The quantitative relationship between pitch and frequency was articulated as far back as Mersenne and Galileo Galilei in the 17th century according to De Cheveigne (2005). Koenig, who invented the tuning fork, also created a tonometric apparatus in the 1860s from which he could demonstrate the existence of different frequencies within a given sound. This development was followed by Helmholtz who described beats of the upper partial tones in Helmholtz (1863). Ellis's (1875) translation of the Helmholtz's (3rd German edition of 1870) *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, further describes a cent system for measuring this, and helped pave the way for intervals to be tempered objectively, creating a scientific basis from which determinations of beats between intervals could be made. Good intonation tries to minimise the frequency of these beats between intervals, which are created as a result of similar frequencies (measured in hertz) of two separate partials belonging to two separate notes 'flanging' together. These pulses or beats, when heard in quick succession, make the interval sound discordant, and produce the sensation of the notes being out of tune.

In this variable, the participant must self-determine the accuracy with which each played note is arrived at, relative to the starting open string. This skill set, in relation to the above, draws on fine motor reflexes cultivated to roll the fingers of the left hand forward and backwards on the fingerboard, to obtain optimum string length/vibration for the note concerned. In addition to the creation of full tones between notes (where fingers are not beside one another) and to the creation of semitones (where the fingers are beside one another), the participant must demonstrate an understanding of the difference in their playing. That is, not having fingers spaced too narrowly to produce tones which are flat, and not having the fingers spaced too far apart, making semitones too sharp. The difficulty experienced by novice players to differentiate between tones and semitones manifests itself when novice players make equal spaces between the fingers, producing tones which are flat, and semitones which are sharp.

Because of this, it is of interest to the observer to examine the extent to which the interval between 1st and 2nd finger – E to F sharp on the D string, for instance – is a full tone. The tendency would be to make this distance homogeneously narrow. That is, to make all of the notes in a Major scale the same distance apart, which they are not. Similarly, for instance, on the A string, the distance between the C sharp and the D natural should have a tight spacing between the 2nd and 3rd finger when in tune, to produce a true semitone interval between the two notes in first position. The order of tones and semitones in a piece of music depends on the scale type (key signature) which the piece belongs to. Major scales, for instance, have a semitone between the 3rd and 4th note, whereas Minor scales have a semitone between the 2nd and 3rd notes. Music which is familiar, is generally constructed from patterns which adhere to either major or minor scale structures.

Variable 2 therefore incorporated a complete D Major scale to monitor the participants' ability to differentiate between tones and semitones, made by the 1st, 2nd and 3rd fingers on both the D and A strings. The first octave only was processed by Melodine software for all participants, to obtain the intonation score. Students who played higher grades, requiring more than one octave scales, also had only the first octave analysed. Measuring upper octaves, where the student must change position in the process, would have diluted the observation question. This is because difficulties in changing position (which will affect intonation higher on the fingerboard towards the bridge) pertain to a separate singularity, relating to shift technique and accuracy. The observation question is adequately determined in the first position, as fundamental flaws in intonation can be detected in first position regardless of playing level.

3.7.2 Related work

Today, this question is being addressed with computer applications such as 'Musicwrench' developed by Doyle (2016). This application, which is similar to Melodine, displays notes as you play them and gives direct feedback relating to the intonation of the notes being played. In Figure 12 below, the intonation feedback (G Major) is displayed, showing the notes played which are in tune on the screen in black. Notes which are flat appear in red, and ones that are sharp appear in blue. The display underneath the music stave indicates the degree to which the notes are sharp, flat or in tune. This application for recording and delivering feedback directly to the learner became available after the research study at hand began. This work indicates a need for direct feedback about intonation during learning.

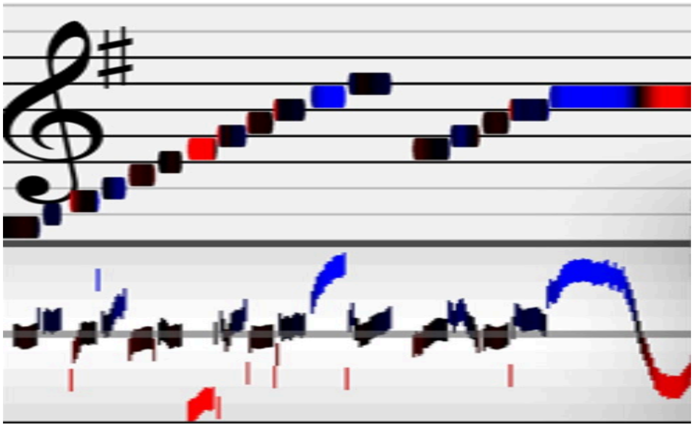


Figure 12. Musicwrench Feedback Doyle (2016)

3.7.3 Recording of data

> "How can music performance be studied scientifically? We base our analysis primarily on information available in the sound alone."
> (Parncutt and MacPherson, 2002, 200)

Violins are tuned in fifths with the second highest string A being tuned to 440 Hertz. Variable 2 calculates the total number of cents 'off' each predetermined frequency, for each fundamental of the given scale rendition of D Major. The results are plotted into a table of incorrect frequencies played, to display the total number of cent outside the fundamental note frequency. This was achieved using visual transcripts of recorded scales produced by Melodine. The software calculates the deviation and records the exact cent quantity off the fundamental note frequency, for each note played in the scale performed (cents are 100^{th} divisions of a semitone).

It is postulated that if discrepancies exist in a student's playing, these can be detected in the sound alone. For this reason, it was felt that a simple scale would be appropriate, as if subtle discrepancies existed between tones and semitones in the playing of the scale, the same discrepancies would exist in a piece of learnt music, such as would be provided in an examination context. It was also noted that students who progress to higher grades can still have intonation issues with simple scales, which is detectable with the technology. In this regard, four such students were asked to play a two octave scale, but only one octave - the lower one, was analysed in the same way as the other students, for the purpose of the intonation variable determination.

In the example below, for student #23 an accumulated score of 330 cent was recorded. Each note played, registered a specific number of cent sharp (sharp or flat will be registered equally in the addition of these discrepancies). Discrepancies of the open A string were subtracted from this figure, in order to eliminate any temperament issues, and ensure that tuning issues would not adversely affect the participant's score, (in this case, 13 and 6 cent, respectively, were eliminated, arriving at a figure of 311 cent total discrepancy). The maximum hypothetical intonation error is given at 1500 cent (that is, 100 cent for each note of a single octave on the D Major scale).

In the below example, for student #23 an accumulated score of 330 cent was recorded. Each note played, registered a specific number of cent sharp (sharp or flat will be registered equally in the addition of these discrepancies). Discrepancies of the open A string were subtracted from this figure, in order to eliminate any temperament issues, and ensure that tuning issues would not adversely affect the participant's score, (in this case, 13 and 6 cent, respectively, were eliminated, arriving at a figure of 311 cent total discrepancy). The maximum hypothetical intonation error is given at 1500 cent (that is, 100 cent for each note of a single octave on the D Major scale).

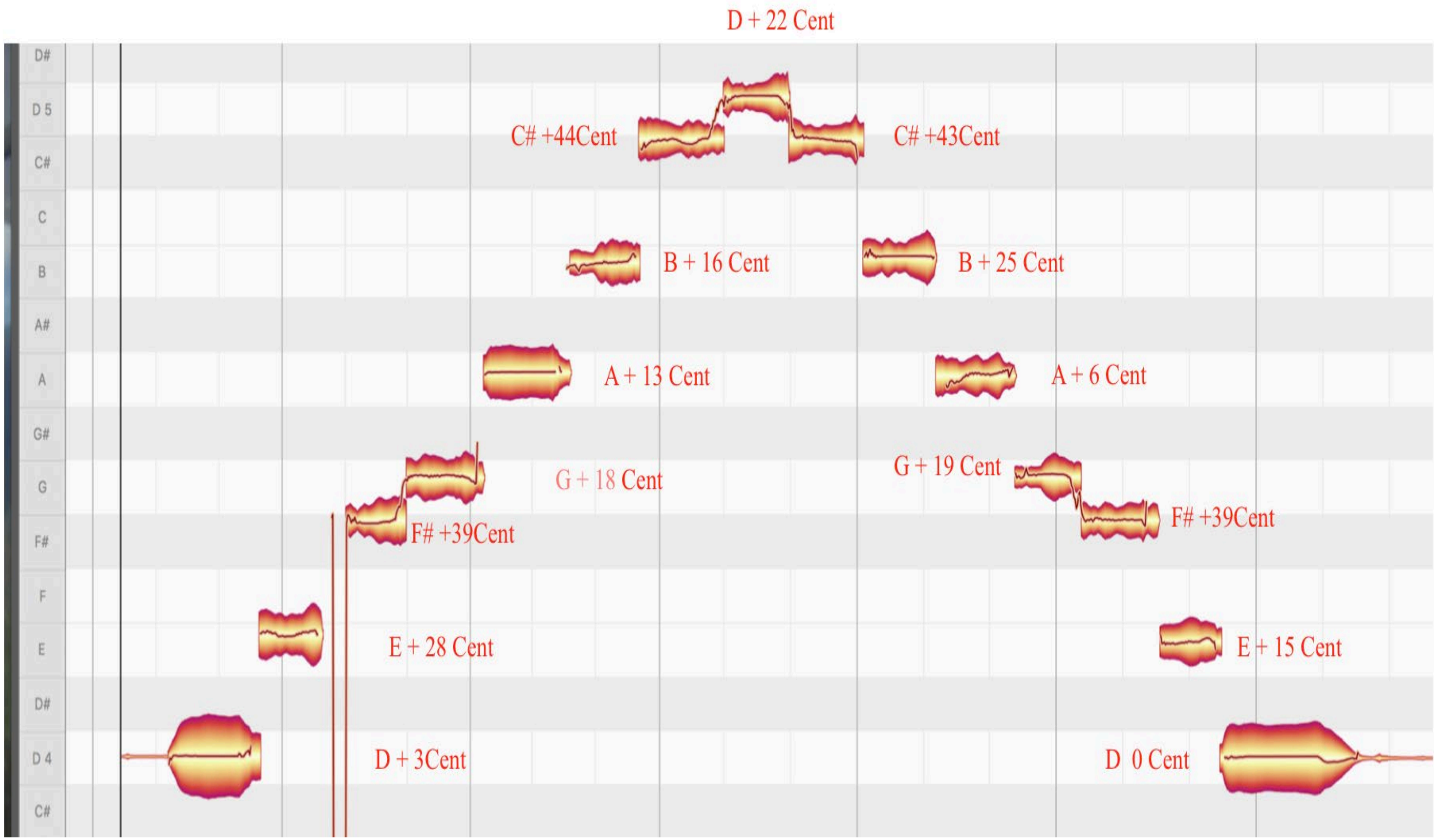


Figure 13. Pitch Transcription.

3.7.4 Calculations involved

The following example provides the path for calculating the intonation variable 2 of the above participant. (See Appendix 3 on page 222). Zero is excellence and 1500 is the worst. The student scored 311 cent scored errors, therefore the student got 1500-311 = 1189 correct. 1189 out of 1500 is equivalent to (1189 x 100) / 1500 = 79%. The best score would have no cent deviation from fundamental to calculate. The maximum number of cent deviation that could hypothetically be present is 1500, i.e. 100 cent for each note of a one octave scale containing 15 notes off. For example, as can be seen in Figure 13, page 88, a total of 311 cent was recorded by this student after subtracting open strings. This means that the degrees of excellence (the aspect that was correct) was 1500 – 311 = 1189 cent correct. This can now be seen in the context of a total excellence percentage over 1500, hence (1189 x 100)/1500 giving student #23 a score of 79% for the intonation variable 2, for comparison with grade outcome. Similarly, student #3 scored a total of 220 cent errors. The degrees of excellence were therefore 1500-220 = (1280 x 100)/1500 resulting in a percentage score for this participant of 85%

3.7.5 Observation protocol

The observation protocol for variable 2 took place again at the school, where the students were taking group lessons on one morning, as an alternative to their regular class, as arranged with the principal. Each time, two students arrived together, and one waited while the other performed the task. Again, if students felt that they could repeat or do a better job, this option was provided for them, as it was not a sight reading test. Most students were happy with their first rendition of the scale. The performance was recorded on a MacBook Pro, with the sound file being sent directly to Melodine software. The files were stored on a hard drive and analysed later, to determine the exact number of cent discrepancies, to arrive at a participant's score for variable 2 – intonation.

3.7.6 Feedback to learner

Using a whiteboard projector, it was possible to display and replay animated versions of Melodine's computer screen outputs of individuals' recordings (see Figure 13 above). It was also possible to pause on given notes, and highlight the exact pitch deviation measured in cent, while listening to the sound (frozen in time). This proved very useful, as many students being new to string learning, were sceptical about critical observation, in regard to their intonation accuracy. From that point of view, it was helpful and reassuring to have specialist comments verified by these machine measurements. In addition, the software was later used to record some of the student pieces, and analysis indicated that it was the same notes which were repeatedly out of tune for a particular player. This supported the assumption that discrepancies in intonation carry forward to pieces, often with the same notes being affected.

The learning that takes place as a result of this feedback helps to broaden out the discussion about what is needed to create and maintain good intonation in other keys also. Using different finger patterns, the same problems in relation to the spacing of fingers can be discussed and overcome. For instance, the semitones which tended to be inaccurate between the 2nd and 3rd finger in the extract for many participants, will tend to occur between other fingers in different keys, and as a result of the feedback experience, this can be anticipated by the students. This approach, in time, addresses in a proactive way, underlying causes of bad intonation and as a process, has an impact on musical outcomes further down the line.

3.8 Variable 3 – Rhythm Accuracy

3.8.1 Theoretical underpinnings

For people to play music together, they must firstly be able to agree on how their endeavours relate to time. The largest time unit in a music system is the piece of music itself. This is broken down into units of phrases – AABA etc. The phrases are reduced into individual bars in each section – bar 53, for instance. The bars are further divided into structures of twos, threes, or fours, etc., with the number of beats per bar notated as 2/4, 3/4, or 4/4, meaning two, three or four crotchets in a bar respectively. The beats themselves are subdivided into separate parts: quavers, semiquavers, and demisemiquavers – the smallest parts. These arrangements or groupings of notes tend to reoccur. In music that we are familiar with, it is reoccurrences which create musical familiarity.

Working with precise moments in time, either through referencing or articulation, is fundamental to what a musician does. In writing, rehearsal and performance, meaning that is laden with rhythmic information, exchanged between composer and performer, or between bandleader and band follower, makes explicit when, and for how long, events should take place. This rhythmic component constitutes half of what is going on in music, as code. The other half is concerned with pitch – the frequency that notes occur at in Hertz, or their altitude.

Rhythm is processed separately from pitch according to Miyamoto (2007). Therefore, assessment should ideally be viewed in terms of a separate singularity. Moog (1976) found rhythmic skills to emerge after one and a half years, with a level of competency expected between three and five years of age according to Sims, (1985). So, it can be expected that young children can comprehend the tasks put before them in the test, following learning and preparation. Two tones are perceived as not sounding at the same moment at 5.6 milliseconds apart for five-year-old children according to Reifinger (2006), suggesting a high potential for accuracy from this age onwards. While seven-year-olds perform better than five-year-olds, adult non-musicians' abilities are not statistically different from seven-year-olds' abilities according to Drake (1993), suggesting that acculturation is not a predetermining factor in rhythm accuracy. Further improvement after seven years is only achieved with training according to Smith et al. (1994).

Precision in the articulation of moments within a bar is postulated to be a key component in musicianship and music making, and the basis on which Variable 3 is constructed. Figure 15, below, shows the printout of the rhythm observation responses for one student. The first line on the chart is a response to a request to articulate the first beat in each bar. Participants differ in their ability to determine this moment. However, this moment itself does not differ in location. It is a constant, and is mapped on the chart which divides the beat in question into tenths, for the purposes of scoring the variable, thus measuring participant rhythmic accuracy.

3.8.2 Related work

Iversen and Patel (2008) used the Beat Alignment Test (BAT) for assessing beat processing and beat synchronisation abilities using the software authoring environment Max/MSP Cycling 74. Participants' 'tap times' were recorded by pressing the spacebar on the computer. The study found a wide range of performance ability in tapping off the beat, and significant difference between trained and untrained participants. To assess the association between perception and production Fujii and Schlaug (2013) used subsets: music tapping, beat saliency, beat interval and beat finding, with a psychophysical adaptive stair-case model to determine perception and production thresholds. Their study found wide distributions of individual abilities to tab in synchrony with a beat, and negative correlations between synchronisation and perception and production thresholds.

3.8.3 Recording of data

The 10 listening tasks required students to identify different parts of a bar and clap their hands at the appropriate moment, i.e. the first beat in the bar, the 3rd beat in the bar etc. The test began, and over a 16 bar duration, a sound file was produced for each of the 10 questions. These sound files were not processed during the observation session, but stored for analysis at a later stage. Analysis would involve assessing the best 4 bar section of each sound file, and marking each bar from 1 to 4. This was repeated for each of the 10 rhythm questions, resulting in a maximum score of 40 for this variable. Figure 14 below lists the rhythmic tasks for variable 3.

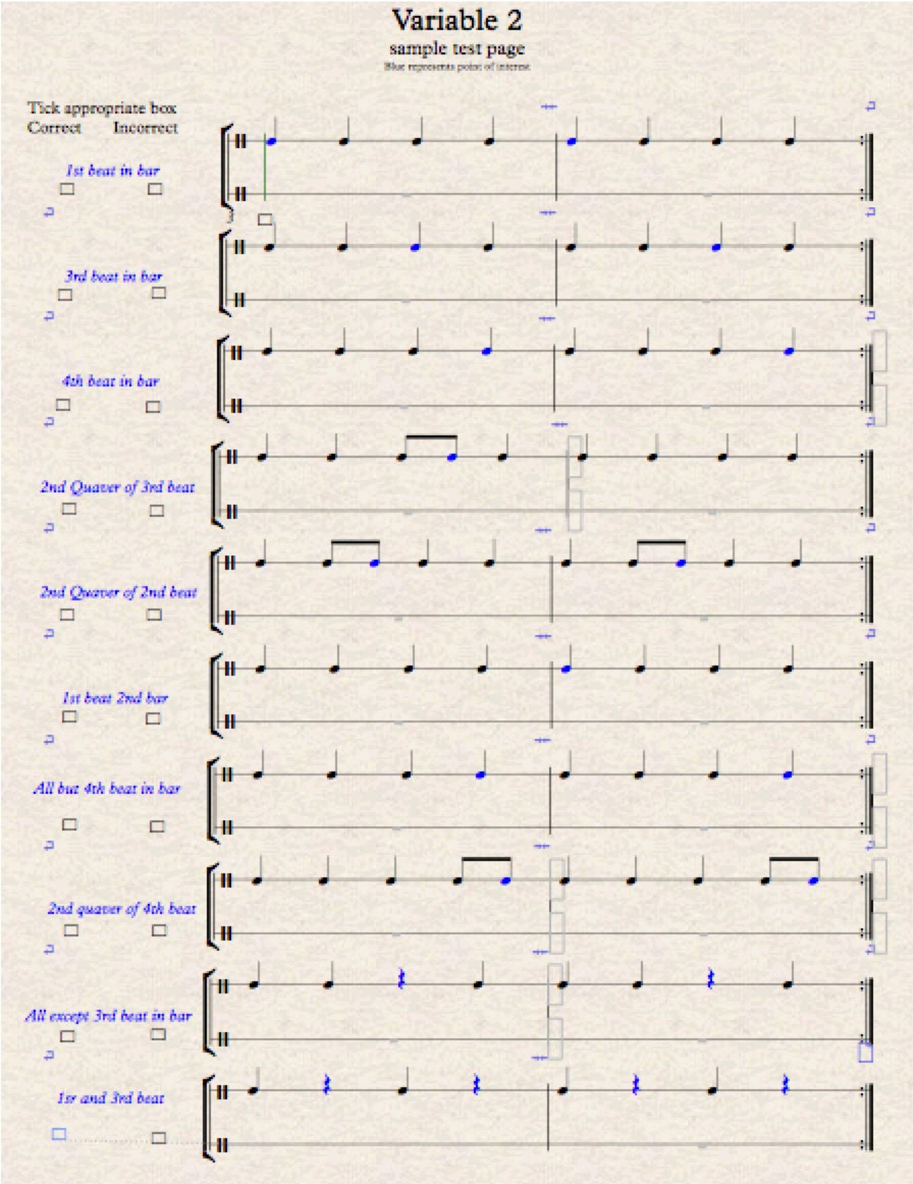


Figure 14. Rhythm Tasks

The 10 listening tasks required students to identify different parts of a bar and clap their hands at the appropriate moment, i.e. the first beat in the bar, the 3rd beat in the bar etc. The test began, and over a 16 bar duration, a sound file was produced for each of the 10 questions. These sound files were not processed during the observation session, but stored for analysis at a later stage. Analysis would involve assessing the best 4 bar section of each sound file, and marking each bar from 1 to 4. This was repeated for each of the 10 rhythm questions, resulting in a maximum score of 40 for this variable. Figure 14 lists the rhythmic tasks for variable 3.

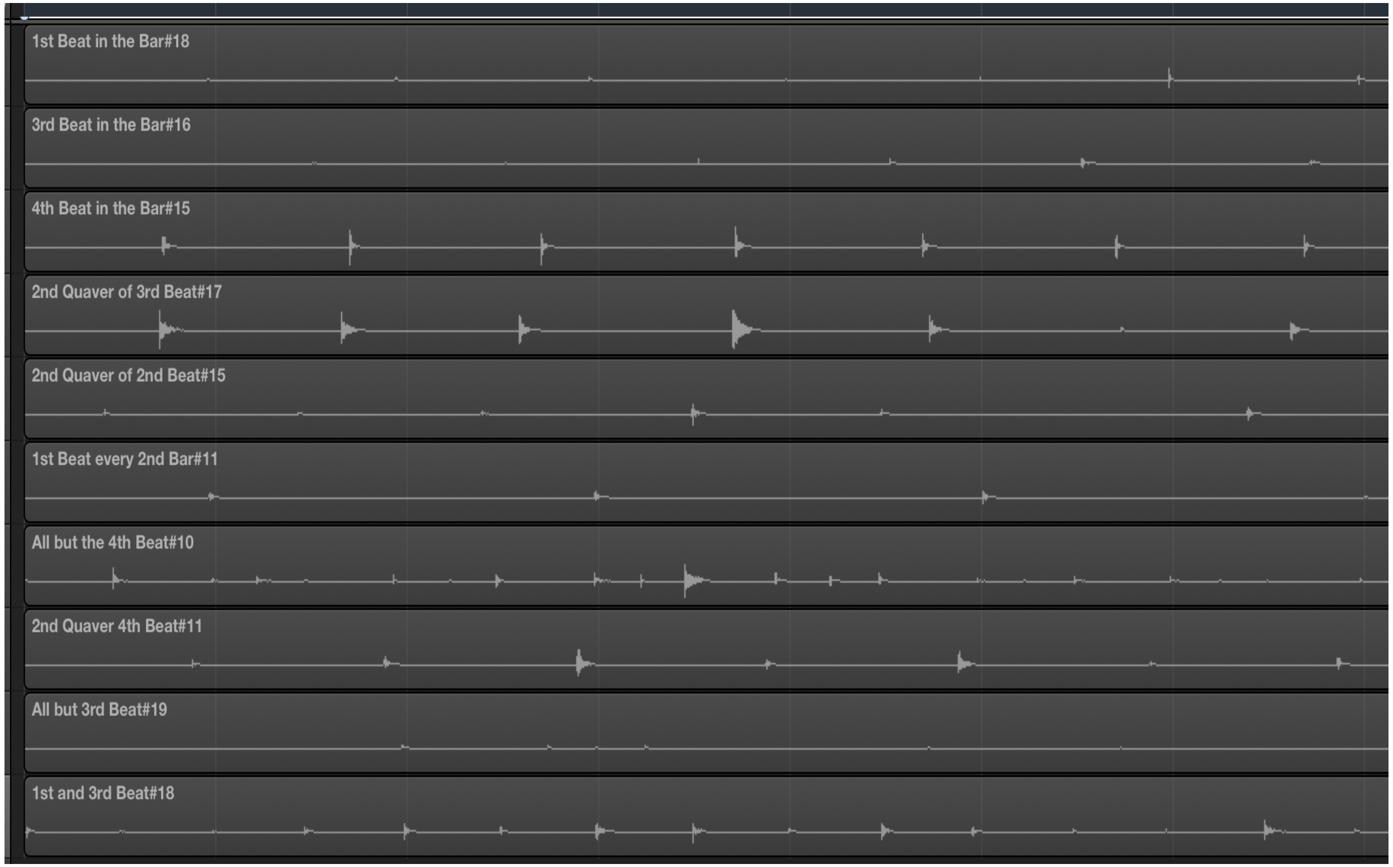


Figure 15. Rhythm Response.

Figure 15 shows a response for 10 questions over 6 bars. For instance, the first beat in the bar (question 1) or the second quaver on the third beat (question 4). The participants are asked to clap their hands on the specified moment in the bar, whilst a maximum of eight bars is being recorded. Using standard recording software (in this case GarageBand was used), the responses are plotted against a timeline in separate layers, which represents each question. Actual sound file responses in the study were often spread over a 16-bar duration.

This visual representation of the recording indicates how responses can be measured. The tolerance of accuracy is set moderately, at less than one quaver. In other words, the participant must clap inside the generous minimum of a quaver within the timeline to score correctly. The participant must identify and respond to four rhythmic moments, for each of the 10 rhythm response tasks – for instance, clap on the first beat in the bar, relative to a consistent metronome beat in 4/4 time provided by the software. A total of 40 correct responses is equal to a 100% score for Variable 3. Analysis consisted of assessing the best 4 bar section of the sound file, and marking each bar from 1 to 4. This was repeated for each of the 10 rhythm questions, resulting in a maximum score of 40 for this variable. The scores are converted into a percentage for comparison with the other variables and grade results.

3.8.4 Calculations involved

The calculation used to arrive at a percentage score for variable 3 relating to rhythm is as follows. 40 x 100/40=100%. So 40 correct clapped responses result in a 100% score. For instance, participant 1 scored 35 out of 40 possible correct answers. Hence we can calculate (35 x 100)/40=87.5% which results in a rounded percentage score for variable 3 of 88% for this participant.

3.8.5 Observation protocol

The participants were prepared for the test in advance, with tutoring on how to go about clapping at certain parts of a bar. They arrived at the observation sessions in groups of two, as before, in the school music room. An Apple MacBook Pro was used, running GarageBand software to record the participants' responses individually. The participants were told to clap their hands at designated points for each of 10 separate metronome listening's. They listened on headphones to a metronome track set to 60 beats per minute throughout, with the main beat in the 4/4 bar being identified with a higher pitch metronome sound. Some students asked for the volume in the headphones to be increased, while others asked for it to be reduced. A pre-test took place, which allowed students to filter out other sounds which may have been taking place in the school ambiance.

Student responses were mixed, with varying levels of intensity of claps being produced, mixed levels of accuracy, and inconsistent maintenance of flow during the observation session. These inconsistencies presented several problems during the data analysis of the observations, notably a difficulty in identifying if the participant understood the task but was unable, in places, to implement it, or if they did not understand the task, but were able to implement it. The lesson learnt from this was to keep to a singularity principle, that is a separation of construct within the rhythm test framework. This would allow for observations of rhythm comprehension in isolation from rhythm implementation, as a level of ambiguity could otherwise exist, when trying to extract this information from the audio files, as was experienced in some cases.

3.8.6 Feedback to learner

Students enjoyed having their rhythm responses played back to them in class. The simplicity of having their hand clapping recorded and played back in class provided great amusement for many of the students, with a certain amount of surprise for others, at the lack of continuity in the hand clapping audio extracts. The more serious aspect of understanding why importance was being placed on ‘places’ in the bar, and developing the ability to articulate them accurately through clapping, were also discussed. Future models of the variable construct would benefit from incorporating ways of measuring comprehension as a separate entity to implementation, as it was unclear during the feedback sessions and data capture process, on which aspect some participants would need to concentrate, in order to benefit from the exercise.

3.9 Variable 4 – Instrument Angle

3.9.1 Theoretical underpinnings

The way the violin is held, discussed in detail by Courvoisier (2006) and others, is central to an effective and sustainable approach illustrated by Fleisch (1939). The shoulder rest method is favoured, as it lends practical support to youngsters struggling to hold the instrument. As a result, cognitive overload, stemming from the multiple competing challenges in string learning becomes minimised, and attention can then be directed at other tasks, such as tone production, intonation and literacy. This transition from supporting the instrument with the left elbow against the trunk of the body (pre-baroque) to independent support with a shoulder rest (post-classical) is, in many ways, evidence of greater ergonomics being employed, as technique evolves. Greater technical demands, with position changes where one plays higher up on the finger board, have necessitated this independence of hold, allowing the left hand to traverse the fingerboard unimpeded by difficulties associated exclusively with holding the instrument, particularly when descending.

3.9.2 Related work

Whilst the literature does not refer exclusively to studies relating to the angles which novice violinists hold their instruments, there is substantial reasoning put forward by tutors explaining why the instrument should be played with strings close to the horizontal plane. The obvious theoretical reasoning being that when strings are almost horizontal, the bow will not require input from the player simply to keep it there. Resting the bow on the string, as it is guided backwards and forwards, produces a better tone than holding the bow in place, whilst moving it across the string. Attempts are made with chin and shoulder rest fixture arrangements, to reduce the incidence of instrument ‘drop’ and make it as easy as possible for students to hold the instrument correctly. While Ramella et al. (1995) recommends that a central chin rest was found to reduce chin lateral deviation, children in the study all used chin rests fitted to the left hand side of the tail piece, with a curved shoulder rest for added support.

3.9.3 Recording of data

The internal 720p HD camera on a MacBook Pro laptop was positioned at right angles to the strings, and at the same height as the player’s instrument, and approximately five meters away from where the player was standing (see examples in Figure 16 below). A short video file was produced for each participant. The files were imported into Kinovea motion capture software and cropped, to measure the best angle predominantly maintained in the extract, relative to a horizontal line on the wall behind where the player was standing. Measurement was made for the angle between strings, and the line parallel to the ground on the wall in the background. In this construct, the more parallel the strings of the instrument are with the ground (the smaller the angle plus or minus), and the higher the variable will be marked.

3.9.4 Calculations involved

The internal 720p HD camera on a MacBook Pro laptop was positioned at right angles to the strings, and at the same height as the player's instrument, and approximately five meters away from where the player was standing (see examples in Figure 16 below). A short video file was produced for each participant. The files were imported into Kinovea motion capture software and cropped, to measure the best angle predominantly maintained in the extract, relative to a horizontal line on the wall behind where the player was standing. Measurement was made for the angle between strings, and the line parallel to the ground on the wall in the background. In this construct, the more parallel the strings of the instrument are with the ground (the smaller the angle plus or minus), and the higher the variable will be marked.

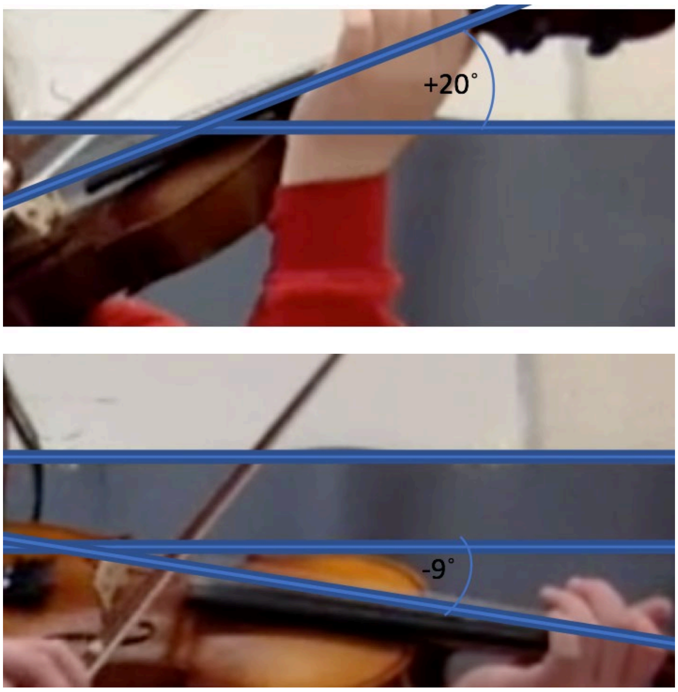


Figure 16. Instrument Angle

3.9.5 Observation protocol

The students were informed that the observation was related to the way they hold their violins and they were encouraged to hold the violin in the way that they had been shown. The students were then invited to play a short piece of their own choice. They generally played one of the pieces being prepared for their grade test, but on occasion some students played other pieces, usually related to what was being learned in school. Each student was informed that a short extract would be recorded, using the internal camera on a MacBook Pro laptop positioned orthogonally from the strings, at the same height as the players instrument, and approximately 5 meters away from where the student was playing. The video extract was retained for the analysis phase of the research. The students enjoyed this observation session, as no other preparation was required. It was seen as an opportunity to simply play a piece of music, with students arguing about who played the best, and generally enjoying the experience. No discussion was undertaken to explain the rationale behind the observation, other than to encourage the students to hold their violins correctly.

3.9.6 Feedback to learner

It is not uncommon to hear a string teacher say 'hold up your violin'. In my experience as a string teacher, I find students are often surprised to discover that, despite thinking they are holding up their violin, in reality they are not. This has consequences over time. The main one being the effect it has on the quality of tone produced, owing to unnecessary pressure or tension exerted to hold the bow in place, rather than simply balancing it. In class, it is easy for fellow students to form an opinion about another player's posture in regard to this, but it is very difficult for the player to form an opinion about their own playing, themselves. The feedback offered here points out, first, if it is an issue, and secondly, to what extent it must be corrected. Students welcomed the idea of having the degree to which they were holding up their violin quantified, which was interpreted by them in general terms as "good", "not so good" or "could be better." The point being that a dialogue about this aspect of posture was initiated for the student, to which they could then respond. It also helped to explain that the more level the strings are with the ground, the less mechanical force that must be exerted with the right hand to hold the bow in place, allowing the bow to glide unimpeded across the string, agitating it mostly through its own weight and velocity.

3.10 Variable 5 – Bow Angles

3.10.1 Theoretical underpinnings

The Helmholtz motion is a phenomena which occurs when a bow is drawn across a string at its maximum potential, when the intersection takes place at 90°. This motion translates directly into a sound quality, which is uniform in shape and consistent in tone. String teachers try to instil in students the need for them to cultivate the ability to draw the bow across the strings, as close to 90°as possible. Gerle (1991), reiterates how the angle between the bow and string at tip, middle and frog should be as close to 90° as possible, ensuring a consistent orthogonal trajectory and minimising skew, or similar characteristics, which compromise tone production.

Variable 5 sets out to measure this, and makes three pertinent assumptions relating to elbow angle. Firstly, when the bow is at the tip, the angle at the elbow (between the ulna/radius and humerus) should be at a maximum (about 160°). Secondly, when the bow is at mid-point, the angle at the elbow should be close to 90°. Thirdly, when the bow is placed at the frog, the angle at the elbow joint should be at a minimum (in the region of 60°). When these criteria are achieved, the trajectory of the bow is most satisfactory. For instance, if the angle at the elbow is still at 120° when the bow is at the frog, movement higher up at the shoulder would have occurred to compensate this, and the trajectory of the bow is, therefore, compromised. In addition to this, by viewing the three orthogonal bow angles, also at point, middle and nut, the observer can corroborate the extent to which the Helmholtz motion will be diminished.
Schoonderwaldt (2010) noted that a straight bow may not be seen by all as an essential prerequisite, however from a novice teaching perspective, it is universally accepted that a straight bow, as a teaching goal, moves the learner in the direction of a correct bowing technique. Gerle (1991) advises that deviations from this norm are permissible for advanced players, to adjust the bow’s distance from the bridge.

3.10.2 Related work

Written in response to criticisms about the advocacy of rotations in bow movements, Hodgson defends the use of photography in bow angle measurement thus:

> “motion photography has enabled me to give more accurate analyses of the basic movements made by all competent performers than has hitherto been possible, and that we should make use of the information gained by this form of research. The first experiment

> shown to a beginner by every intelligent teacher, irrespective of school, demonstrates the fact that bowing in a line approximately parallel to the bridge produces a finer tone than so called 'crooked' bowing." (Hodgson, 1935, 347-348)

Rotations encourage fluency over stop-start motions, but can easily be confused with a bowing which is not 'straight'. The use of photography was incorporated to point to the benefits of a correct bow trajectory. This is still the case today. More up to date projects, like the '3d Augmented Mirror', indicates how the problem of bow trajectory can be tackled in an innovative way, by using today's technology.

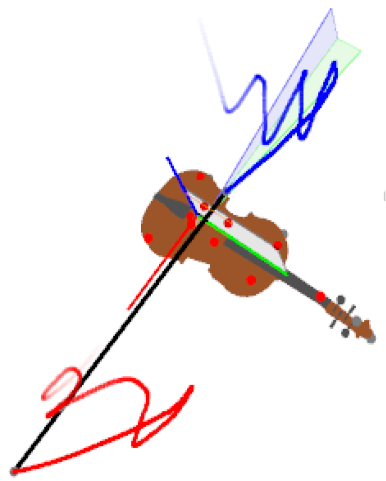


Figure 17. An example of data collected by an Augmented Mirror (Ng et al., 2007)

The *i*-maestro 3D Augmented Mirror project (as shown above) above, was developed at Leeds University by Ng et al., (2007) and consists of a 12 camera infra-red motion capture setup. AMIR interfaces with a VICON bridge system over the Ethernet using TCP/IP streaming 3D co-ordinates of markers to Max MSP at 200fps. The system provides posture feedback, and develops postural self-awareness during practice. This interactive multimedia environment creates a new pedagogical paradigm, which maximises self-learning possibilities. Technology used in this way provides a solution to the problems which arise during teaching, and helps to measure the level of progress being made. However, the amount of laboratory support needed to implement this approach made it an unsuitable choice for the study, even though its objectives and aspirations are very similar and directly relate to the research topic.

Elements in the *i*-maestro 3D Augmented Mirror project reoccur in this study, such as determinations about orthogonal bow angles and bow distance. With less detail, but sufficient accuracy to make clear determinations about angles and bow distance, Kinovea software was used in the analysis of video recordings to measure the same criteria, and provided the data for variable 5 (Trajectory) and variable 6 (Bow Distance) (see Figure 18, on page 104, and Figure 19.1-2 on page 109). This approach dispensed with the need for a laboratory, and made field observation of these characteristics a reality.

Inertial motion capture systems (Synertial) differ from optical motion capture systems, insofar as they do not rely on multiple cameras to gather data. An inherent problem with camera use in data collection, aside from the restricted use due to ethical considerations, is the problem of occlusion. Occlusion occurs when markers on the subject being observed become obscured by the movement of the subject during the observation. While this problem can be overcome by incorporating multiple cameras, triangulation of this data from the multiple sources makes for a cumbersome approach to field work. Inertial motion capture systems obtain data by placing sensors on the body of the player directly, to record rotational movement and dispense with the need for a dedicated studio space. Another advantage of the inertial motion capture is the possibility of the system collecting the data prior to being hooked up to the host computer. Kinovea software provided a software solution which enabled elbow and orthogonal bow angles to be measured with high levels of accuracy. This would enable bow and orthogonal angle determinations to be made for variable 5.

Kinovea is a video player software package developed originally for sports enthusiasts to help monitor movement and posture. It has many facilities which enable the user to slow down, compare, track and evaluate movement. It has been adopted in the research study, because of its intuitive interface and its transferability to the tasks at hand in the observance of novice string player movements. It was used to track bow movement, to measure distance covered, determine bow angles at strategic points in the bow trajectory cycle, and observe bow direction movements, in relation to sight-reading test recordings. Tools in the software enable the user to superimpose angle measurement and tracking features onto video clips for analysis and comparison. There is nothing in the literature to suggest that this software has been used previously in observation studies of novice musicians. Hence the approach adopted in this study is original, and represents a novel addition to the body of knowledge on this topic.

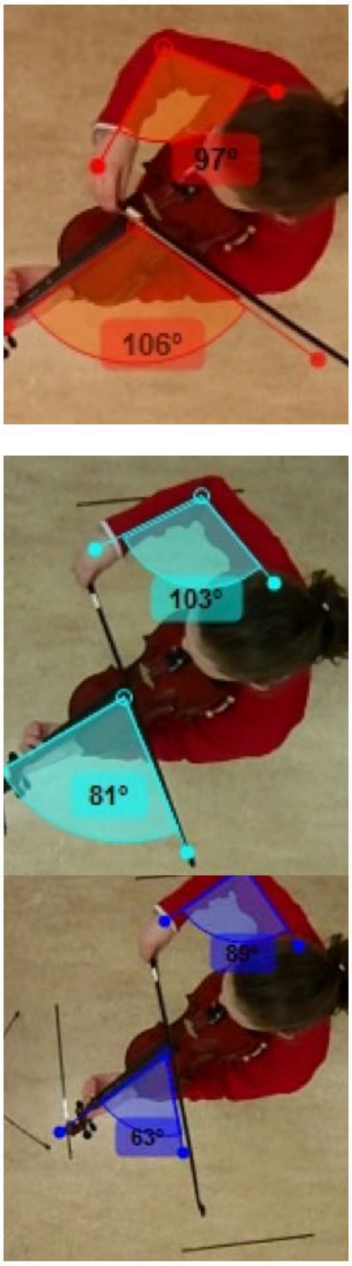
3.10.3 Recording of data

The bow angles data was recorded at the same observation session as the instrument angles data. In order to record the data needed for the variable 5 construct, a camera was also placed above the music stand, in the music room where the student was to perform, at a distance of about 3 meters. This created an overview of the students' bow movements, and using Kinovea software made it possible to make accurate measurements about three separate elbow angles, at the tip, at the middle and at the frog of the bow, and the three orthogonal bow/string angles also at the point, mid and tip of the bow. These six angles would later be combined to create a single percentage bow angle value for variable 5, relating to bow trajectory. All angles were recorded as 'best possible', that is, the closest to point, middle and frog, as demonstrated in the video extracts. The observation is concerned with the three elbow and three orthogonal bow angles observable during performance at these extremities. Hence, on the third illustration of Figure 18 below, the player feels she is playing at the tip of the bow, when in fact she is not, hence a score of 56% for this aspect of bow trajectory, is determined by a small elbow angle at the point.

Variable 5

Code 1

Elbow and Orthogonal
bow angle component analysis



Optimum Angle	Angle Recorded	Offset	Degrees of Perfection	Efficacy
Elbow 60°	At Frog 97°	-37°	23°	38%
Orthogonal 90°	Orthogonal 106°	-16°	74°	82%
90°	At Mid 103°	-13°	77°	86%
90°	Orthogonal 81°	-9°	81°	90%
160°	At Pt 89°	+71°	89°	56%
90°	Orthogonal 63°	-27°	63°	70%

422% /6

Variable 5 Total **70%**

Figure 18. Elbow and Orthogonal Angles

3.10.4 Calculations involved

The following explains how the angles were calculated, to arrive at a percentage score for variable 5. The variable consists of six components, three relating to the right elbow angle and three relating to the orthogonal angle between bow and string. The calculation approaches the problem of measurement firstly by defining what is described as degrees of perfection, by subtracting the error from optimum angle. Hence the absolute differences between the ideal and the actual values are converted into percentages. So if 60 degrees is perfect, and a student demonstrates an angle of 97, then they are 37 (i.e. 97-60) degrees from perfect. Similarly, if a student demonstrates an angle of 23, then they are also (23-60) = 37 degrees from perfect.

The first component of nut elbow angle is calculated as follows. Normalising the scores into percentages, the prefect score of 60 degrees will be equal to 100%, therefore 1 degree is the equivalent of 100/60 = 1.67%. Therefore 60 degrees is perfect, but we need to express the level of perfection as a % out of 100. If a student scores 97 degrees s/he is 37 degrees from perfection (i.e. 37 degrees from 60). Therefore, measuring the degree of accuracy, rather than error, is calculated as the perfect score (60) less the degree of imperfection (37) = 23. Normalising this score, gives 23 x 1.66 = 38.318, so we are measuring the extent of excellence for the first component of the bow trajectory variable 5. For clarity, the calculations used in Figure 18 are given on page 226.

Whilst the bow is at the nut, the orthogonal bow angle is also measured. The calculation for this measurement is 90 degrees equals perfection, i.e. 100%. Normalising the scores into percentages, therefore 1 degree is 100/90 =1.11. If a student demonstrates an orthogonal angle of 106 degrees at the nut, this is a deviation of 16 degrees from a perfect score. This equates to 90 – 16 = 74 degrees of excellence. Converting this score into percentages, is calculated as: 74 x 1.11, which gives a value of 82% for the second component of variable 5.

The calculations needed for the 3rd, 4th and 6th components of the trajectory variable use the same mode of mathematical calculation as in each of these cases, where deviation away (either above or below) from the desired optimum 90 degrees angle is measured. For example, looking at the 3rd component calculation (that of mid elbow), if a student demonstrates a mid elbow angle of 103 degrees, this is a deviation of 13 degrees from a perfect score, and hence can be calculated and concerted into a % value, as follows. Thus, a student displaying 77 degrees of perfection, will receive a percentage score of 77x1.111 = 85.54% rounded to 86%

Continuing with this example, the 4th component calculation is of the mid orthogonal angle where 90 degrees is perfect (and equivalent to 100%, where 1 degree = 1.11%), such that a student displaying an angle of 81 degrees is 9 degrees offset, but has 81 degrees of perfection, which is equivalent to a percentage score of 89.99% rounded to 90% (i.e. 81 x 1.11). Likewise, for the 6th component, the point orthogonal angle, 90 degrees is perfect (and equivalent to 100%). If a student displays an angle of 63 degrees, this is 27 degrees offset but has (90 - 27) = 63 degrees of perfection, which converts in to a percentage score of 69.99%, rounded to 70% (i.e. 63 x 1.11)

The remaining 5th component is the point elbow angle, which uses the same calculation, but with a larger perfect angle of 160 degrees. As 160 degrees is a perfect score, this is equivalent to 100%, therefore 1 degree is the equivalent of 100 / 160 = .625%. If a student displays a point elbow angle of 89 degrees, this is 71 degrees from perfect, and hence 89 degrees of perfection, which converts to a % score of 55.62% (i.e. 89 x .625), rounded to 56%.

Having obtained the 6 component variable percentage scores, they are added together and divided by 6 to arrive at an average trajectory score. Hence in the examples calculated above, we have, Nut elbow angle = 38%; Nut orthogonal angle = 82%; mid elbow angle = 86%; mid orthogonal angle = 90%; point elbow angle = 56%; point orthogonal angle = 70%. These total to 422, and the mean value = 422% / 6 = 70%

3.10.5 Observation protocols

The observations took place simultaneously with the variable 4 session, but using a separate 640 x 680 resolution camera at 30 fpm mounted overhead at a distance of 2 meters. Similarly to observation 4, students arriving in the music room were generally excited about the observation session, as they could perform a short piece of their own choice. Most students performed their 'best' exam piece, but some chose to play a scale or something they had learned from memory outside the class. The point of the observation was to gather enough information to determine (from a best case scenario) the bow movements in relation to elbow and orthogonal angles for each student, when playing at the middle and each extremity of the bow. The observations took place during class for school students, and after class for private students, and determinations were made at a later stage, based on both video clips gathered.

3.10.6 Feedback to learner

The tendency to control the bow high up on the arm, with movement between the humerus and the shoulder, is discouraged in string teaching. The lower down the movement takes place, the more likely it is for the optimum trajectory angles, mentioned above, to be achieved. Again, this approach opened up a dialogue using visual evidence to support the goal of the teacher in encouraging the student to improve bow trajectory, by observing these angles. Students were encouraged to alter their posture by the visual evidence in the short clips, which demonstrated the lack of movement at the elbow joint causing the problem of compromised elbow and orthogonal angles. The observations here also highlighted the obstacles to playing with a full bow, which is the subject of variable 6.

3.11 Variable 6 – Bow Distance

3.11.1 Theoretical underpinnings

This construct is concerned with the amount of bow used by the participant to execute a selected observation passage. Violin teachers frequently request the student to use more bow. Looking objectively at this phenomenon, it is now possible to measure, with a high degree of accuracy, the distance travelled by the bow, and calculate this as a definitive nominal quantity, making it clear how much bow was used, relative to any given ideal amount. The total amount of bow used to play the extract (a single octave scale) is measured in centimetres, quantifying the participant usage and populating the raw data fields for variable 6 (bow distance).

It is postulated, in accordance with bowing theory by Galamian (1966) and Gerle (1991) that the more bow used, the greater the tone production potential. This is supported by Helmholtz motion theory, which was discussed on page 100. While the bow angles will impact directly on the quality of sound produced, the amount of bow used in relation to the Helmholtz motion will affect the consistency of the tone produced and the dynamic potential of the sounds being made.

The amount of bow used to produce a given tone increases the likelihood of the note being clear, consistent and less susceptible to splitting, having blemishes or disintegrating. A challenge in string teaching is to encourage novice string players to maximise their use of the bow. This is achieved primarily through an optimum 'opening' of the right elbow joint, to produce the greatest coverage at all three stages of nut, mid and point, discussed in Variable 5. Often, a restricted opening of the elbow joint is incorrectly compensated for, through excessive shoulder movements and faulty bow trajectory.

3.11.2 Related work

Other studies which address the problem of students not using enough bow include the MusicJacket by van der Linden et al. (2011a). The jacket, which was designed to give vibrotactile feedback to correct novice string player bow trajectories, found feedback could also be delivered when the student had used the optimum amount of bow during field studies. This discovery proved successful in achieving the goal of promoting maximum bow use. The 3d augmented mirror by Ng et al. (2007), although not developed for the field, was also helpful in directing string learners' attention to bow usage deficits, through visual computer generated simulations of student movements, which were then compared to model bow movements.

3.11.3 Recording of data

Students were requested to play a scale of D Major, consisting of 15 notes, using maximum bow possible. The amount of bow used was then measured with Kinovea software's object tracking capability. The length of the bow was multiplied by the notes played (15 in number) to find the optimum distance value for the extract, taking into account 2 bow sizes and the necessary calculations to match them to the performance of the extract. Only one octave was recorded, both ascending and descending, to make the determination about the amount of bow used by the student to play the scale.

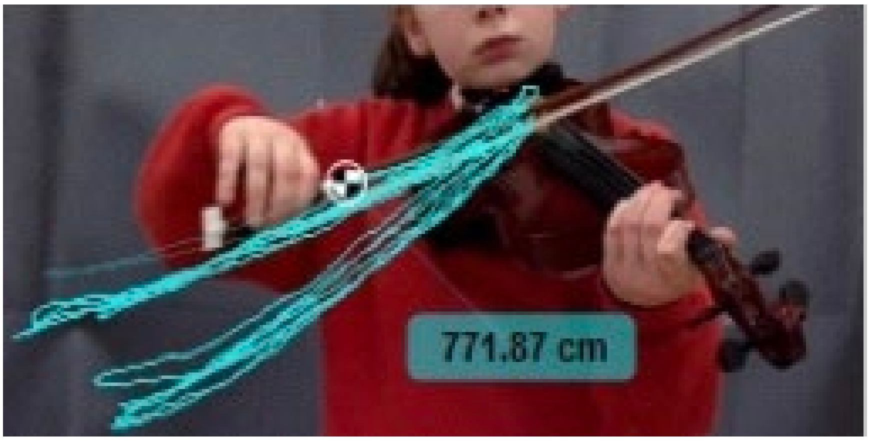


Figure 19.1 Bow Distance Tracking ½ size

Figure 19.2 Bow Distance Tracking ¾ size

Figure 19.1-2 illustrates how a ‘fix’ is made on the bow; a small white sticker is attached to the bow which can be tracked when in front of a black background. Kinovea software can then follow that point of movement throughout the duration of the observation. This fix enables the distance travelled by the bow to be quantified by the software, given that the length of the bow is already known. The resulting machine measurement figure, given in centimetres, constitutes the distance travelled by the participant’s bow whilst playing the scale, thus defining variable 6 - bow distance measurements.

3.11.4 Calculations

The normalisation calculations calibrate the bow length with optimum bow distance travel for 15 notes. The ½ size maximum bow length of 62 cms is expressed for 15, giving a maximum distance of 930 cms of bow travel if the extract is played perfectly.

As with the other previous variables the total possible score of 930 cms of bow travel is converted to percentages, where 1 cm = 100 / 930 – 0.107%. If a student utilises 772 cms (out of 930 cms) to play the extract, then the normalised score is 772 x 0.107 = 82.6%. Where the violin is ¾ size rather than ½ size, then the maximum bow length is 68.7 cms, but the mode of calculation is the same. Hence the maximum possible distance for bow length travel for 15 notes, played perfectly will be 15 x 68.7 = 1030, which converts to 100%, (where 1 cm = 100 / 1030 = 0.097%). If a student utilises a total of 583 cms out of a maximum of 1030 cms, then the normalised score will be 538 x 0.097 = 52.28%

3.11.5 Observation protocol

On this occasion, the camera in the MacBook Pro was used to record students playing a single octave D Major scale. The camera was placed so as to observe the bow length throughout the extract performance (adjacent to the travel path of the bow), at about 4 meters from where the student was standing (as shown in Figure 19.1 and 19.2). A fix (a small white label attached to the nut of the bow) enabled Kinovea software to track the movement of the bow in the analysis stage, as this fix contrasted with the black background mounted behind where the student stood. On occasion the tracking 'jumped' off the sticker during analysis and it was necessary to re-run the analysis (usually at a slower replay speed) to ensure that the software followed the movement of the bow fix continually, from the initial first note of the scale onset, to the break in contact between the bow and string on completion of the final note, where a machine measurement in centimetres was recorded by the software for bow distance travelled.

3.11.6 Feedback to learner

Kinovea motion capture software enables the bow to be tracked, with distance covered being calculated and displayed in centimetres as the sequence is played. This is later compared to the optimum distance required to play the extract. For many novice string players, the bow does not stray very far from the middle and it was viewed with interest by the students, when it was pointed out (with footage extracts run through the Kinovea software), that the bow was in fact in the middle, when the students thought that they were playing with full bows as requested.

3.12 Variable 7 - Guidance

3.12.1 Theoretical underpinnings

One of the first obstacles which a novice string player must overcome, is the ability to play the intended string note without also playing the adjacent strings unintentionally. The bow's path must move on the same plane to avoid hitting adjacent strings. When playing at the tip of the bow, there is a wide margin of error available, but when playing at the frog, this margin decreases substantially, making it more likely to inadvertently play another string. Another problem occurs during the directional change of the bow. Reversing the direction is best achieved with a continuous flattened loop action. If the flatness of the loop is not maintained, however, the potential for the bow touching adjacent strings increases. Overcoming these problems requires accurate motor skills in relation to: bow placement; bow plane orientation; bow trajectory action; and wrist rotation. This array of tasks needed to isolate single tones can be overwhelming for the novice player. The sound of more than one string being played, while trivial, is isolated as a singularity, and a predictor for this learning milestone - the focus of variable 7.

3.12.2 Related work

There is no known research work which assesses this particular trait as a singularity, and, while much has been written about the complexity of double stopping - where the student purposefully learns to play more than one string at the same time, little is known about monitoring novice players, as they struggle to isolate single strings. Elemental tasks feature less in the literature than expert ones, yet many hours are spent by the teacher trying to help novice players overcome this hurdle. The cellist Pablo Casals is known to have spent many hours playing open strings. Little has been written about measuring this relatively minor achievement, as it is taken as a given in string playing, and the process of achieving mastery of it is so far removed from the product (the performance) which receives most of the attention. However, without mastery of the muscle memory needed to produce unblemished tones, flaws consisting of unwanted string sounds can be detected at any stage of an extract with audio machine measurement software, such as Melodine.

3.12.3 Recording of data

This observation took place during the variable 8 – sight-reading session. Using the same video footage, the observer measured the bow plane singularity by counting the number of times that more than one note was played (double stopped) during the 20 note sight-reading extract. Half of the participants scored 95% or higher, so the variable could also have been conceived as a binary output to register a positive or a negative mastery of the problem. It was felt however that counting the number of double stops would give a finer grained indication of ability. The scores were converted into a percentage score from 20 degrees of perfection.

3.12.4 Calculations involved

Perfection is defined as no double stops (playing more than one note) when playing 20 notes. Hence 0 errors = 20 degrees of perfection, equivalent to a normalised score of 100%, where 1 error equals a 5% reduction in the score. For example, if a candidate gets 4 errors, this results in 16 degrees of perfection which represents a score of 80% for variable 7, as 16 x 5 = 80%.

3.12.5 Observation protocol

Observations took place during school hours for class students, and after school hours for private students. The observation protocol was the same as for the reading observation protocol, given that both observations stemmed from a single video footage extract. It was pointed out to participants at the outset that they should focus on playing single strings. The fact that the main body of attention was directed towards the sight reading element caused attention to focus on that aspect, with individual strings becoming a secondary consideration. However, it was pointed out that students should play as close to the bridge as possible, to minimise the likelihood of playing more than one string.

3.12.6 Feedback to learner

The feedback sessions took place before and after school hours for both groups (classroom and private students) about two months after the data was gathered. The main discovery for students in these sessions centred around the knowledge that they were more inclined to play more than one string if the bow had travelled onto the fingerboard. It was evident from the video clips that incidences of more than one string being played was caused by a combination of the bow being in the fingerboard area where the effects of the curve of the bridge is at its minimum (making it harder to single out inside strings), and poor control of the single plane on which the bow travels. Outside strings were understandably least affected by this. The separating out of individual tasks, such as playing single strings only, helps to direct attention and reduce the incidence of students feeling overwhelmed by the combination of tasks which must be undertaken simultaneously.

3.13 Variable 8 – Reading Ability

3.13.1 Theoretical underpinnings

In the experience of the observer, students tend to overlook the importance of simple sight-reading tasks. As the student moves to higher grades, basic sight-reading gaps in learning become masked. While context specific considerations described by Mitchell and Green, (1978) impact on overall efficacy, reading quality does not always keep pace with other developments, and reading gaps in reading tests overlook this. It is doubtless the case that the context, including (for example) the key chosen, the complexity of a time structure, the familiarity with the genre, or anxiety levels present, will all influence the musical outcome of a sight-reading test. However, by measuring singularities, rather than testing for shortcomings which are more closely related to musicianship inadequacies than reading, variables 8-10 probe the main sub-divisions which constitute a sight-reading task. In this design, it is the notes' pitch, the notes' durations and bowing indications which are considered. Like the other variables, the sight-reading test is looking at gaps in comprehension which can occur at any level. Mishra (2014) found sight-reading improves as the musicality of the performer improves and not simply by attending to the visio-motor decoding process. Lehmann and McArthur (2002) suggest:

> "it is the gap between each person's ordinary level of rehearsed performance and the same person's ability to perform at first sight that is the problem. The smaller the gap is, the better the sight reader". (Lehmann and McArthur, 2002, 136)

While good sight-reading cannot solve problems of technique, it bridges a difference gap between first reading and polished performance. The physical properties of the stimulus when reading, described as bottom-up, and how the reading task matches the expectations, described as top-down, combine to determine the reader's overall skill. As the reader becomes more skilled, a preoccupation with performance can obscure learning gaps which may have occurred along the way.

Set between a grade 1 and 2 standard for all participants, the observation construct reflects basic abilities which can be difficult to detect, as tests become more complex, as causes of inaccuracy become multifaceted. To preserve validity, overall quality of playing, which often has an important bearing on sight-reading outcomes, was not considered in this test. The ability to play the correct notes written on the page, is the first consideration. Secondly, accuracy of note durations, which reflect simplistic cohesiveness of the rhythmic structure, were observed and thirdly, the accuracy with which the student acts upon the bow indications were considered.

The measurement criteria are separated out and located within the 20 note extract (as shown in Figure 20), corresponding to 20 degrees of perfection for each sub division of the observation – variable 8 durations, variable 9 durations and variable 10 bowing indications respectively.

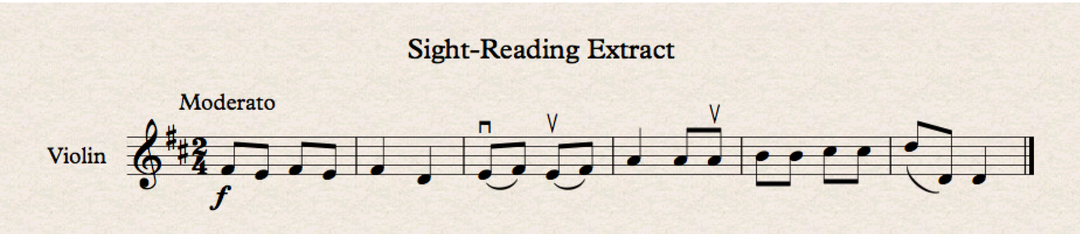


Figure 20. Sight-Reading Extract

Notes should be played accurately and in tune, with attention being paid to F and C sharps; that is, with finger patterns facilitating the 2nd finger to be beside the 3rd. The correct durations, in which crotchet and quaver differentiations should be clear and explicit, with the participant demonstrating a clear understanding of the difference between the two. Bowing indications at bar 3, 4 and 6 should be observed. Interpretation, musicianship, quality of playing and other musical outcome considerations are not taken into account in this test, in order to maintain focus on the aforementioned elements of novice ability.

3.13.2 Related work

There is evidence of similar singularity driven observational studies but without isolation from a musical context. Figure 21.1 below illustrates a web based rubric from *RCampus* (a forum for creating rubrics under diverse learning and assessment headings) located under the heading of Music Sight Reading Performance Assessment. The individual components which mark progress in reading singularities are explicit, but to a large extent obscured, because of a focus on the product or musical outcome. Looking through the criteria in this rubric, one encounters some headings unrelated to reading which is meant to be the focus of the test design. Instead, the rubric observes items such as intonation and pitch accuracy. This diminishes the ability of this type of reading test, to identify where the actual reading gaps exist, as accuracy is likely to have more to do with motor skills than reading comprehension.

In the second example rubric taken from *RCampus* (Figure 21.2) it is easier to extract information about reading ability as the singularities are more explicit and accuracy is not a consideration. The errors are quantified more clearly and the test is likely to be applied by the examiners more evenly as a result of a simplified criteria structure.

	No Concept 1 pts	Beginning 2 pts	Developing 3 pts	Advancing 4 pts	Accomplished 5 pts
Rhythm Accuracy	No Concept Extremely inaccurate, difficult to count number of mistakes, as if notes are played randomly out of time	Beginning Mostly inaccurate, most rhythms are played incorrectly, but there appears to be a minimal amount of understanding of the notation	Developing Mostly accurate, more than two mistakes and mistakes have noticeable impact on the performance	Advancing Nearly perfect, one or two mistakes, but mistakes do not detract from the performance	Accomplished Perfect, no noticeable mistakes
Pulse	No Concept No perceptible pulse, as if notes are played randomly	Beginning Very inconsistent pulse, but rhythms still seem potentially related to A pulse, even though the pulse changes or starts/stops	Developing Somewhat inconsistent pulse, many noticeable fluctuations that detract from performance	Advancing Consistent pulse almost always, with only a few musically inappropriate fluctuations	Accomplished Consistent pulse throughout, with musically appropriate fluctuations, but no fluctuations that detract from performance
Pitch Accuracy	No Concept Extremely inaccurate, difficult to count number of mistakes, melodic contour is non-existent	Beginning Mostly inaccurate, more mistakes than correct pitches, even melodic contour suffers	Developing Mostly accurate, more than two mistakes (some of which are blatantly incorrect fingerings), but melodic contour is still present	Advancing Nearly perfect, one or two mistakes (mistakes are likely a result of a missed partial, not a blatantly incorrect fingering)	Accomplished Perfect, no obvious mistakes, all pitches are accurate
Relative Intonation	No Concept Almost every note seems drastically out of tune with the next/previous, difficult to distinguish an overall pitch center	Beginning Most notes are obviously out of tune (untrained listener would easily notice)	Developing Two or more obvious problems (untrained listener would easily notice)	Advancing One or two slight, but noticeable problems (untrained listener might NOT notice)	Accomplished No noticeable problems with intonation
Articulation	No Concept Nearly all articulations are ambiguous, as if the way in which notes are attacked/released is not even considered by the performer	Beginning Most articulations are ambiguous or blatantly incorrect	Developing Good, more articulations are correct/unambiguous than not.	Advancing Very Good, almost all articulations are clear with only a few ambiguous or incorrect articulations	Accomplished Excellent, all articulations are clear and correct

Figure 21.1 Sight Reading Rubric (a)

Dynamics	No Concept No perceptible and intentional change in volume	Beginning Very limited range of dynamic contrast AND multiple mistakes in dynamic accuracy	Developing Limited but perceptible range of dynamic contrast OR multiple mistakes in dynamic accuracy	Advancing Slightly limited range of dynamic contrast OR only minor mistakes in dynamic accuracy	Accomplished Wide range of dynamic contrast and all dynamics are executed clearly and correctly
Style	No Concept No discernible style and no obvious attempt to play in a certain style	Beginning Unclear whether style was considered, style is ambiguous, but potentially appropriate aspects are present	Developing Obvious attempt at style, but with major inconsistencies, inappropriate aspects, or inaccuracies	Advancing Obvious attempt at style, but with minor inconsistencies, inappropriate aspects, or inaccuracies	Accomplished Clear, appropriate, and successful representation of style, could be used as an example of this style
Tone Quality	No Concept Tone is unclear, inconsistent, very unpleasant to listen to, and is not recognizable as this instrument	Beginning Tone is very inconsistent with major deficiencies in clarity and resonance; it is not representative of the instrument, but is still recognizable as this instrument	Developing Tone is inconsistent, particularly between different ranges, with obvious deficiencies in clarity and resonance; it borders between appealing and unappealing	Advancing Tone is generally consistent across all ranges, but with minor deficiencies in clarity and resonance, but still representative of the instrument and pleasant to listen to	Accomplished Tone is consistently clear, appealing, resonant, and highly representative of the instrument across all ranges; could be used as an ideal example of this instrument
Musicianship	No Concept No perceptible attempt to play with any "musicality", no "musical" decisions seem to have been made	Beginning Unclear whether aspects of musicality were considered, but potentially some "musical" decisions were made and executed poorly	Developing Obvious attempt at musicianship, but with major weaknesses in execution of musical decisions	Advancing Obvious attempt at musicianship, but with minor weaknesses in execution of musical decisions	Accomplished Successful and appropriate musical performance, could be used as an example of a "musical" performance

Figure 21.1 Sight Reading Rubric (a) continued

	Poor 1 pts	Fair 2 pts	Good 3 pts	Proficient 4 pts	Advanced 5 pts
Rhythm Reading	Poor Missed rhythms in more than 9 measures	Fair Only missed rhythms in 7-9 measures	Good Only missed rhythms in 3-6 measures	Proficient Only missed rhythms in 2 measures	Advanced Played all of the rhythms in each measure correctly.
Pitch Reading	Poor Missed over 13 pitches	Fair Only missed 10-13 pitches	Good Only missed 6-9 pitches	Proficient Only missed 3-5 pitches	Advanced Played all of the pitches in each measure correctly.
Dynamic Reading	Poor Missed appropriate dynamic levels in more than 9 measures	Fair Only missed appropriate dynamic levels in 7-9 measures	Good Only missed appropriate dynamic levels in 3-6 measures	Proficient Only missed appropriate dynamic levels in 2 measures	Advanced Correctly incorporated all of the dynamics marked in each measure.
Articulation Reading	Poor Missed appropriate articulations in more than 9 measures	Fair Only missed appropriate articulations in 7-9 measures	Good Only missed appropriate articulations in 3-6 measures	Proficient Only missed appropriate articulations in 2 measures	Advanced Correctly incorporated all of the articulations marked in each measure.
Steady Tempo	Poor Steady tempo was not maintained throughout most of the measures	Fair Steady tempo was maintained throughout except 7-9 measures	Good Steady tempo was maintained throughout except 3-6 measures	Proficient Steady tempo was maintained throughout except 2 measures	Advanced Maintained a steady tempo throughout each measure.

Figure 21.2 Sight Reading Rubric (b) (rcampus)

3.13.3 Recording of data

The reading construct limits itself to three separate variable headings. These singularities include reading notes- variable 8, reading durations - variable 9, and reading bow indications - variable 10. Recording the data involved observing the short video recording of a 20 note sight-reading passage which was played by the student one time, after briefly looking over the piece. The transcript of the recording was analysed at a later stage to determine if the correct notes were played, if the notes were played for the correct duration, and if the correct bowings were used. Scores for the 3 variables - 8, 9 and 10 were then added together to arrive at a total reading percentage score.

3.13.4 Calculations involved

The calculations involved in all of the reading variables use the same equation as for the guidance variable 7. The same extract is used and the same criteria adopted to measure reading ability, based on the performance of the extract.

Perfection is defined as 0 errors over 20 elements, which is the equivalent to a percentage score of 100%, where 1 error equals a 5% reduction in the score. So, if a candidate had a score of 3 incorrect notes relating to variable 8, this would be equal to 17 degrees of perfection, and the percentage score ascribed to variable 8 would be (17 x 100) / 20 = 85%. The same mode of calculation is used for variable 9, where a score of 2 incorrect durations relating to variable 9 would be equal to 18 degrees of perfection, which would give a score for variable 9 of (18 x 100) / 20 = 90%. Similarly, for variable 10 if a candidate had a score of 3 incorrectly observed bow directions, this would be equal to 17 degrees of perfection, and so the score ascribed to variable 10 would be calculated as (17 x 100) / 20 = 85%. The three percentage scores are added together and divided by 3 to calculate the mean which is the final value for the reading percentage variable. Hence in this instance, the final value is 85 + 90 + 85) / 3 = 87% total reading score.

3.13.5 Observation protocols

The students arrived in the music room in groups of two, and were presented with alternate sight-reading extracts for the reading observation. There was a proportion of students who were less enthusiastic about the reading test, as there is a culture of learning how to play traditional music from memory, and students who excel at this are somewhat sceptical about the usefulness of reading notation. However, preparations were made by all students in advance, to ensure the students could undertake the test. The test was pitched between a grade 1 and grade 2 reading ability. A selected 20-note passage was given to all the participants, who had the opportunity to study the extract briefly before playing it. The piece included enough elements to enable determinations to be made about all 3 sight-reading criteria mentioned already. Students did not hear the previous renditions of other students. Observations were recorded on a MacBook Pro computer camera and the video footage was later analysed, to make determinations about pitch, durations, and bowing criteria for each participant.

3.13.6 Feedback to learner

Feedback from sight reading tests tend to lack specific detail about how outcomes can be improved. Students who struggle with reading are often unable to create sufficient musical meaning from the sight reading extract, because of a gap in one or more of the elements. The feedback comments associated with conventional sight-reading tests, as illustrated above, and in the section on comparative instructional feedback (see page 156), record aspects of musicianship which may be missing, but obscure elemental root causes of difficulty in reading music. Without clear and quantifiable indications of the strengths or weakness of these elements in the rendition, it is not possible to improve sight-reading ability in a tangible way, based on feedback given.

The separation of the 3 singularities in the reading test ensured that these learning milestones are not obscured by musical considerations. Students in the study were reassured when presented with factual information about their reading ability. Facts, like students' abilities to play the correct notes, were evident, but the amount of time spent on quavers was not correct. Alternatively, the notes and rhythm that were played in the extract were correct but students overlooked the bow indications in bar 4. This type of feedback while excluding important information about related music ability is potentially more effective as a sight reading assessment model. Feedback sheets, which cite comments about intonation or style when they are supposed to be addressing reading ability, overlook cognition and implementation and tell us little about the sight reading issues.

The feedback sessions with students sometimes included going over the video extracts which produced lots of different types of reactions, with statements like 'why do I have to play that bowing' or 'I don't understand how the key signature changes the notes'? or 'I forgot to stop the bow in bar 4 between the quavers and ended up playing one long crotchet instead.' This type of dialog suggests that the students are addressing fundamental issues which help them understand issues which lie at the heart of a music sight-reading challenge.

Asmus (1999, 22) describes measurement as "the assignment of a numeric value that characterises a particular attribute of interest", and states that rules which define how attributes are categorised can bring consistency to the observation. The 34 points of inquiry in Figure 22 below are reduced to eight variables: pitch, intonation, rhythm, instrument, bow total, distance, guidance and reading total, for comparison with grade outcomes in standard ABRSM performance examinations, rated in marks out of 150, but which are also standardised by being converted to percentages.

3.14 ABRSM Grade

Reference to the method of calculation of the ABRSM has been made in section 2.7 on page 49. However, having provide detailed information on the source and mode of calculation of the other variables, for the purposes of exposition before commencing the data analysis, it is worth recapping briefly, the salient details of the ABRSM grading system. The overall mark for an ABSRM grade is awarded out of 150, which is made up as follows: For each of three pieces played, students are awarded up to a maximum 30 marks. For both scales/arpeggios presented, and their rendition of a sight reading test, students are awarded up to a maximum of 21 marks, Finally, for aural responses to four extracts played by the examiner on the piano, students are awarded up to a maximum of 18 marks. Hence, the total score is derived as: score for Piece 1 (30 marks max.); score for Piece 2 (30 marks max.); score for Piece 3 (30 marks max.); score for Scales/arpeggios (21 marks max.); score for Sight-reading (21 marks max.); score for Aural tests (18 marks max.) = Overall Score (150 marks max.). For the purposes of this study, the scores were then normalised out of 100%, such that (for example) a student's ABRSM grade of 75 out of 150 would be converted to 50%.

3.15 Conclusion

The methodologies described above illustrate how novice performance attributes can be determined and quantified as data. The resulting profile represents an objective assessment for comparison with subjective grade outcomes. An epistemological stance rooted in the European art music tradition has been stated in relation to the research approach, supporting the design strategy taken. Ethical considerations have been listed and participant involvement and consent forms have been described. Other perspectives relating to objectivity have been considered, and reference has been made to the research environment, in both the classroom and private context. Changes made following the pilot study, along with improvements to the research design, have been discussed. The operationalisation of the study, conceptual frameworks and theory behind the variable constructs have been explained. The observation schedule has been outlined and the variables' content and calculation have been made explicit. Figure 22 below provides a summary sheet, showing the number and description of all of the variables, whose modes of calculation have been described in detail above, both for the raw scores and also for the conversions into percentages.

The following chapter explains in detail the process by which the data shown in the table above were analysed, the statistical techniques used and inferences drawn, using Pearson's correlation, Linear Regression, and Causal Path Analysis.

Variable Number	Variable Data Exemplars							Variable Description
NA	31	32	33	34	35	36	37	Participant ID
1	5	4	5	7	6	9	6	Pitch Raw Score
	50	40	50	70	60	90	60	Pitch converted to %
2	315	178	238	52	189	80	121	Intonation Raw Score
	70	88	84	97	87	95	92	Intonation converted to %
3	30	23	30	36	31	38	33	Rhythm Raw Score
	75	58	75	90	78	95	83	Rhythm converted to %
4	15	20	15	5	5	5	30	Instrument Angle Raw Score
	75	67	75	92	92	92	50	Instrument Angle converted to %
5a	103	98	83	68	72	91	95	Nut Elbow Angle Raw Score
	28	37	61	86	80	48	42	Nut Elbow Angle converted to %
5b	94	97	93	90	91	90	100	Nut Orthogonal Angle Raw Score
	96	92	97	100	99	100	89	Nut Orthogonal Angle converted to %
5c	98	97	93	91	92	90	102	Mid Elbow Angle Raw Score
	91	92	97	99	98	100	87	Mid Elbow Angle converted to %
5d	95	98	95	90	93	89	79	Mid Orthogonal Angle Raw Score
	94	91	94	100	97	99	88	Mid Orthogonal Angle converted to %
5e	111	108	117	125	121	128	100	Point Elbow Angle Raw Score
	69	68	73	78	76	80	63	Point Elbow Angle converted to %
5f	103	107	106	91	94	92	102	Point Orthogonal Angle Raw Score
	86	81	82	99	96	98	87	Point Orthogonal Angle converted to %
Mean a,b,c,d,e,f	77	77	84	94	91	87	76	Bow Trajectory Mean Value in %
6	698	654	730	800	765	800	645	Bow Distance Travelled Raw Score
	75	63	78	86	82	78	69	Bow Distance Travelled converted to %
7	1	2	1	0	0	0	2	Bow Guidance Errors Raw Score
	95	90	95	100	100	100	90	Bow Guidance Errors converted to %
8	1	1	0	0	0	0	1	Reading Notes Errors Raw Score
	95	95	100	100	100	100	95	Reading Notes Errors converted to %
9	1	2	0	0	1	0	3	Reading Duration Errors Raw Score
	95	90	100	100	95	100	85	Reading Duration Errors converted to %
10	1	2	1	0	0	0	4	Reading Bowing Errors Raw Score
	95	90	95	100	100	100	80	Reading Bowing Errors converted to %
Mean 8+9+10	95	92	98	100	98	100	87	Reading Mean Score in %
11	110	124	102	121	125	120	102	ABRSM Grade Mark Awarded
	73	83	68	81	83	80	68	ABRSM Grade in %

Figure 22. Data gathered, and Variables Calculated and Exemplar values.

4 Chapter Four Data Analysis

4.1 Introduction

By way of recapitulation, the main purpose of the study is an examination of the ABRSM grading systems (which are subjective in nature), through the lens of the ten variables which are shown in Figure 22 above. These variables have been derived from an array of objective physical data on string playing, which has been gathered through the use of cameras and other means of technological data capture. This visual and auditory data has then been analysed, by means of specialist software (Melodine, Garageband, and Kinovea), to provide numerical values for a number a critical features of violin playing, such as (for example) the angle at which the instrument is played, relative to horizontal. The raw data values for these variables have then been converted into percentages. The ABRSM scores used are actually those awarded to the students in the study, and these have also been converted into percentages.

As has been shown, where the ABRSM grading system is used to assess instrument playing and the examiner is a qualified teacher in the instrument being played (piano, violin, cello, etc.), the process involves an assessment of the student's technical prowess along with an assessment of the musicality of the performance. However, the ABRSM frequently uses generalist examiners who may be unable to assess the technical merit of a pupil's playing, and instead based their grading on the musicality of the performance only, as (for example) when an examiner who is a skilled pianist is called upon to assess the performance of a pupil playing the cello. In this chapter, methods of data analysis are used which enable the technical competence of string players to be calculated for ten critical measures which are determined by data captured by means of sophisticated hard and software. These measures, which cover an array of technical aspects of violin playing, are then measured against the ABRSM grade (briefly described at the end of the previous chapter) which the pupils achieved.

Having described the characteristics of the raw data, and the process by which it was gathered, and mathematically converted into % variable values, the next part of the thesis is devoted to analysing the variables statistically. Three types of statistical tests are used. First, Pearson's correlation coefficient is calculated for the degree of association between the independent variables (as shown in Figure 22) and the dependent variable (the ABRSM grade), and the results are assessed at the 0.05% significance level. Secondly, linear regression techniques are utilised, as described in Draper & Smith (1966), and Berendsen (2011), and regression equations, using least squares criteria, are developed to predict the value of the dependent variable (i.e. the grade outcome in the form of the normalised ABRSM percentage scores) from the independent variables (pitch score, intonation, etc.). Finally, following on from the linear regression, causal path analysis is applied, to assess the impact of key independent variables on the dependent grade outcome variable. Visual representations are given, to show how different independent variables have an impact on the dependent grade variable. A summary outlines what can be deduced from the relationships found between the independent variables and the dependent ABSRM variable.

The analysis has extracted meaning from numerical representations of the observations undertaken, by ensuring that the data collected is valid and technically reliable. By this, it is meant that the data accurately reflects measurements taken from observable indicators, deemed to be conducive to good string playing. In addition, regressional models have been developed and finessed to arrive at ones which best fit the data that has been gathered. The creation and examination of these models has been greatly enhanced by powerful computer programs, including SPSS, which have become standard practice in modern statistical analysis.

The objective of the analysis was to examine the data collected from the observation sessions and grade examination results, and determine the degree of association between the two by using correlation, and other methods. The analysis would support or refute the hypothesis statement which postulated a link between objective observation scores and subjective ABRSM grades. Furthermore, exploration and clarification was undertaken, using correlation, linear regression, path analysis and graphical representations of the relationships between the different variables. Correlation coefficients were examined in a correlation matrix to find the strength of relationship between eight independent variables and one dependent variable with those below a 0.05% significance being rejected. The same data was then used to calculate regression lines to determine the predictive power of the independent variables using the least squares criteria. Causal path analysis techniques were incorporated to explore intuitive, inclusive and parsimonious models that could best explain relationships. Causal path modelling was used to probe the strength of effect between exogenous variables, and to isolate those combinations of variables which generated the strongest influence on the endogenous grade variable. Graphical representations of variable score outcomes were drawn to illustrate the relationships in context with regard to student formative feedback. These representations were generated to create a path, through which the findings of the analysis could be returned to the learning environment, and be applied directly to assist student and teacher learning, the hope and aspiration being to create a sense of musical Bildung for both.

In trying to ensure that the variables accurately measured what they had set out to measure, some safeguards were built into the research design. Features such as a singularity of question, dynamic use of technologies, and a separation of overlapping concepts like pitch detection from intonation were incorporated. Also, in the reading variables, a separation of components such as notes, durations and bow directions helped to isolate precise points of enquiry about reading ability. The order in which the observations took place also helped to verify results and maximise observation efficiency. Every effort was made to ensure that participants fully understood the tasks of each observation, and that participants were matched adequately to each task and did not, therefore, suffer any emotional distress or physical discomfort while performing. As mentioned, components which seem to overlap or measure the same thing, can have totally separate cognitive, cerebral or motor functions. Within the pitch discrimination construct for instance, a passive discrimination is engaged, while with the intonation inquiry, participant involvement is discriminatory and active, in order to make adjustments of note frequencies whilst they are being made.

Pupil violinists who display a difficulty in one area of observation can mask problems which stem from a separate technical difficulty. For instance, restricted bow movements were shown in the pilot study to harbour intonation defects, many of which themselves were rooted in pitch discrimination issues. Another example of the need to fully understand what is being measured, was in relation to sight-reading ability, which was made up of rhythm and tonal components. In addition, technical deficits in playing ability can, in turn, obscure which of these aspects of reading difficulty predominates. Consideration was given to the design of the variables, in order to extrapolate from the data, exactly where the root causes of difficulties with musical learning and instrument mastery lie. The potential for the observation sessions to provide validated formative feedback emerged as a result of these deliberations, and became an aspect of discovery and a point of focus for further research.

The reduction of the six bow angle variables (nut elbow angle, nut orthogonal angle, mid elbow angle, mid orthogonal angle, point elbow angle, point orthogonal angle) into one overall score, and the three reading variables into another overall score, also reduced the total number of variables to eight for comparison with grade result. Whilst being quantitative in design throughout, the richness of the variable data collected mitigated against any interpretive component relating to order and direction of associations, the aim having been to determine if there was a relationship between the variables and the grade in the initial hypothesis.

4.2 Data Collection Process

The data collection took place in accordance with the methodology described in Chapter 3. The observation schedule illustrated in Figure 9 on page 76 gave rise to the determination of 34 separate points of inquiry, which are listed in Figure 22 on page 123. The three bow angle measurements and the three right elbow angle determinations were amalgamated into one total bow trajectory (Variable 5). This was done to simplify the process of extracting a single trajectory score from related aspects of the same observation construct. Causal Path analysis later found point elbow angles and nut elbow angles (two of the six bow angle determinations) to be more parsimonious in their impact on grade result, than the other components of independent Variable 5. The three reading variable were also amalgamated into one mean reading variable (Variable 8). All of the other constructs remained as separate determinants of the attributes being observed. In all, a total of eight variables were prepared for comparison with the ABRSM grade results.

Thirty-seven students participated in the grade examinations, out of an initial sample of eighty students who had participated in the objective observation sessions. Many of the school children were reluctant to undertake grade examinations, even though they were capable of passing the examination comfortably. This reluctance can, in part, be explained by the fact that the general population of classroom students in Ireland do not take a music examination at primary level in school. Cost may also have been a factor, but the most likely explanation is that there is not a culture of taking practical music examinations at primary level in Ireland. Music performance examinations take place outside the normal primary school curriculum framework privately, and musical instruction is not encouraged in the same way as, for example, sports activities. Students also may have felt under pressure to undertake the grade examination as part of their participation in the research study. To minimise the possibility of any pupils being put under any stress, by feeling obligated to undertake the ABRSM examination, the sample was therefore reduced to 37 participants who had completed all five observation sessions *and* a grade examination.

4.3 Stages in statistical analysis

First, descriptive statistics are calculated indicating the range, mean, and standard deviation. Then, Pearson's Product Moment Correlation coefficients are presented. Following this, regression line equations are calculated using least squares criteria. Finally, path analysis is conducted with Intuitive, Inclusive and Parsimonious models. Throughout these analyses, graphic charts are displayed to show relative scores for each participant on each variable. Variables: pitch accuracy, intonation, rhythm accuracy, instrument angle, bow trajectory, bow distance, bow guidance and reading ability are compared with grade result. The grade data on which the Pearson's correlation value is modelled is the awarded ABRSM mark (out of 150) reduced to a percentage score out of 1000 for comparison and computation.

In Chapter 1 Section 5 on page 8, the null hypothesis stated as follows: "Relationships between objective observation scores and subjective grade results are random, HO $p=0.5$". The alternative hypothesis asserted that there is a link between the two, therefore H1 $p\neq0.5$. Data collected from objective observation and subjective grade results and analysed at the .05% significance level has demonstrated a significant relationship between variable scores and grade outcomes. Inferential statistics indicate the likelihood of an experimental hypothesis being accepted or rejected.

Variable	N	Range	Mean	Std. Deviation
Grade	37	28	78.55	6.933
Pitch	37	60	64.86	16.935
Intonation	37	19	89.23	5.968
Rhythm	37	85	77.38	16.127
Instrument	37	67	75.54	15.780
Trajectory	37	26	84.44	6.636
Distance	37	43	72.31	11.449
Guidance	37	35	93.38	8.979
Reading	37	27	94.04	7.059

Figure 23. Descriptive Statistics

We could hypothesise that if the independent variables have a relationship with the dependent variables, we would expect that the values for the means and standard deviations would be similar. In fact, the mean value for the ABRSM grade variable is 78.55, and this figure is exceeded by the mean values of five of the independent variables, but in three instances, the mean value of the grade variable, exceeds those of three of the independent variables. Similarly, we might anticipate that the spread of variation around the mean would be similar for our independent and dependent variables. In fact, this is not the case. The standard deviations, representing the amount of variance between the means of each variable, are given in Figure 23 above. None of the mean values are below 60 and two are above 90. Here we can see that for the N=37 participants, the greatest variance occurred within the pitch, rhythm and instrument variables. The Pitch variable had the lowest average score but the largest standard deviation. The standard deviations for these variables are the highest, which means that these variables hold least support for the alternative hypothesis. In contrast, variables including intonation, bow trajectory, sight-reading and bow guidance registered between $\sigma = 5.968$ and $\sigma = 8.979$. In contrast, the ABRSM grade results had a standard deviation of $\sigma = 6.933$.

4.4 Pearson's Correlation Coefficient

Pearson's correlation coefficient is a widely used statistic to describe the degree of association between two variables. Following frequent use of this statistic across a range of empirical analyses, the following bands provide some indication as to the strength of Pearson's r. If Pearson's r is 0.7, then r^2: = 0.7 x 0.7 = 0.49, that is nearly half of the variation in one variable, is explained by variation in another.

High correlation: 0.5 to 1.0 or -0.5 to -1.0.
Medium correlation: 0.3 to .05 or -0.3 to -0.5.
Low correlation: 0.1 to 0.3 or -0.1 to -0.3.

Where the correlation coefficients are positive, this indicates a direct relationship between two variables and where the coefficients are negative, this indicates an inverse relationship. Two tailed tests are used when the positive or negative nature of the relationship between two variables is not known; one tailed tests are used when we know the likely nature of the relationship, but are wishing to test the strength of the relationship. Hence, the two-tailed test explores whether the relationship is significantly greater or significantly less, whereas a single tailed test only explores one or the other. For this reason, the two-tailed test was chosen. Correlations have been carried out and assessed to see if they are significant at the 0.05 level. The significance level for all the tests in the thesis was set at 5%, however, where the significance of a statistical test is greater than 5% (for example, where it is 1%), this has been reported. The decision to use a 5% (as opposed to a 1% or 10% significance level) is arbitrary but, as has been pointed out, Gall et al. (2007) and Cowles and Davis (1982) report that a 5% significance level is invariably used in studies of this kind, and across the social sciences.

	Pitch	Intonation	Rhythm	Instrument	Trajectory	Distance	Guidance	Reading
Pearson's R	.337*	.349*	.103	.598**	.487**	.382*	.537**	.609**

*=significant at 5% **=significant at 1%

Figure 24. Pearson's Correlations between ABRSM Grade and the Independent Variables

As can be seen in Figure 24, the highest correlation coefficient of r=0.609 was recorded for the relationship between the reading variable and ABRSM grade result, and it indicates that it is very likely that reading abilities have a significant impact on performance ability. The reading ability variable was constructed by measuring abilities in the three areas of music reading for strings: reading notes, reading durations, and reading bow indications. From this we can infer that developing music reading skills increases the potential for achieving a higher performance grade result. In only one instance, that of rhythm, there was not a relationship with ABRSM grade which was significant at the 5% level. One possible explanation for this could be the inability of the rhythm test construct to adequately separate cerebral from motor ability. That is, to determine if the participants understood the questions but were unable to articulate the precise execution of the tasks or if the participants were uncertain about the task questions but would have been able to carry them out had they understood them more fully. A better understanding of this would possibly lead to a more representative data set for variable four- Rhythm.

The next most notable correlation occurred between the variables instrument angle and grade result, at $r = .598$ (which is significant at the 1% level). Instrument angle relates to the extent to which the strings are parallel with the ground. As discussed in the section Chapter 3.8 above, relating to this variable, strings should ideally be parallel to the ground, to facilitate free movement of the bow without exerting additional control to prevent the bow from moving towards the fingerboard (instrument too low), or towards the bridge (instrument too high). This uninhibited right-hand technique therefore contributes to tone production and a consistent Helmholtz motion. The result indicates a significant positive relationship between instrument angle and grade outcome.

In relation to variable 7, a correlation of $r = .537$ (significant at the 1% level) indicated how the ability to guide the bow across an individual string, without inadvertently touching other strings, correlated to overall grade result. The tendency to introduce extraneous sounds in addition to intended ones, as discussed in Chapter 3.11 above constitutes a component of tone production which is isolated in this variable. The adverse effect of this tendency comes as no surprise to a trained teacher of string playing, and it was reassuring to see this isolated construct identified at the early stages of observation to correlate so significantly in overall grade results.

The next most significant correlation was between bow trajectory and grade outcome. This variable was the product of six individual bow angle determinations which amalgamated to produce an overall bow trajectory score. The individual components, as described in the section on Variable 5, discussed in Chapter 3.9 above, pertained to three orthogonal angles of the bow as they cross the strings, at point, mid, and frog (both ends of the bow and the middle). The other three angles related to right elbow angles illustrated in Figure 18 on page 104. Unsurprisingly, the resultant 'trajectory' variable proved to be significantly correlated to grade result, at r = .486.

Analyses of inter-relations between individual variables indicated even higher levels of correlation shown in Figure 25 below. A coefficient of r = .861 (significant at the 1% level) was recorded for the correlation between bow guidance and reading ability. While good bow guidance assists the implementation of what has been read in a music score, good reading skills also makes it easier to concentrate on playing the correct string.

The strength of relationship between instrument angle and reading ability was r = .766 (significant at 1% level). This was the second highest correlation and counter-intuitively indicated a connection between reading accuracy and posture. To the trained eye, however, this relationship is consistent with a good technique of holding the instrument independent of the left hand. This predisposes the player to fix their gaze on the music score, without being distracted by reflexes, such as looking at fingers to try and correct finger dexterity inaccuracies. Finger inaccuracies and subsequent intonation faults are often caused by incorrect posture where the palm of the left hand is incorporated to support the holding of the instrument, disadvantaging best finger placement technique, achieved when holding the instrument correctly and horizontal. Again, one aspect has the effect of influencing the other.

		Grade	Pitch	Intonation	Rhythm	Instrument	Trajectory	Distance	Guidance	Reading
Grade	*r* coefficient	1								
	Sig. Level	NA								
Pitch	*r* coefficient	.337	1							
	Sig. Level	.041	NA							
Intonation	*r* coefficient	.349	.381	1						
	Sig. Level	.034	.020	NA						
Rhythm	*r* coefficient	Not Signif	.383	Not Signif	1					
	Sig. Level		.019		NA					
Instrument	*r* coefficient	.598**	Not Signif	.338	.510**	1				
	Sig. Level	.000		.041	.001	NA				
Trajectory	*r* coefficient	.486**	Not Signif	.534**	.393*	.741**	1			
	Sig. Level	.002		.001	.016	.000	NA			
Distance	*r* coefficient	.382	Not Signif	.410	Not Signif	Not Signif	Not Signif	1		
	Sig. Level	.020		.012				NA		
Guidance	*r* coefficient	.537**	.327	.455**	.527**	.578{**	.712**	.450**	1	
	Sig. Level	.001	.048	.005	.001	.000	.000	.005	NA	
Reading	*r* coefficient	.609**	.348	.382	.574**	.766**	.708**	.415*	.861**	1
	Sig. Level	.000	.035	.020	.000	.000	.000	.011	.000	NA

Only correlations that are significant at the 0.05 level of significance for a two tailed test (or above) are included in the table.
**Correlation is significant at the 0.01 level (2-tailed).

Figure 25. Pearson's Correlation Matrix

-inhibiting factors like bow hair contact with strings and restricted bow movement, mentioned earlier. The regression analysis (see below) clearly supports this finding.

The correlation between the guidance and trajectory variables came in at r=.712 (which is significant at the 1% level). This is not surprising considering that both variables are essentially dealing with the same question, albeit from different points of focus. The Guidance variable looks at the extent to which more than one string is played at any given time and trajectory, which quantifies an accumulation of three bow and three elbow angles. If the orthogonal bow angle is compromised, it is likely that the bow will come in contact with more than one string. This becomes more pronounced the further away from the arch of bridge the bow is applied to the string. The relationship between guidance and trajectory was anticipated to be significant.

The r = .708 correlation between reading and trajectory, although significant at the 1% level, is a less obvious one. The question as to whether good technique precedes a display of good reading skills, or whether the execution of good reading skills is made possible by a secure technique, has been left largely unanswered, that is, the order in which these variables might interact. The important point from the analysis has been that there is a significant relationship between the two variables, and that improvements in one are reflected by improvements in the other.

The guidance and instrument angle correlation of r = .578 (significant at 1%) is, to a large extent, self-explanatory. Guidance, relating to the control of which string is played, is extrinsically linked to the angle of the instrument in relation to the ground. While violinists may move the instrument freely up and down whilst playing, particularly in moments of high drama and intensity, further analysis of this phenomenon, however, reveals that the whole upper torso also moves in line with the angle of the instrument, to compensate the divergence. The problem, from a pedagogical point of view, arises when the novice player moves or drops the instrument angle without adjusting the upper torso, creating a change in bow trajectory across the strings. This variable is particularly useful in directing attention to a common cause of bow trajectory divergence, thus confirming the value of this approach in providing meaningful feedback to students on their performances.

A high correlation of r = .574 (significant at 1%) is to be expected between reading and rhythm, given that half of the information to be deciphered relates directly to rhythm, and the other half to pitch. Having said that, it should be noted that a low rhythm score and a high reading score, or visa-versa, directs focus to the component of reading ability which is most in need of attention. Reading ability should not be confused with performance ability. Lehmann and McArthur (2002) point out that it is the gap between the level of rehearsed performance, and an ability to perform at first sight, which becomes narrower as sight-reading ability improves in Parncutt and McPherson (2002). Having examined the inter-relationship of the variables by means of correlations, the analysis now moves on to use a more powerful statistical tool, linear regression.

4.5 Linear Regression

Linear regression enables the prediction of a dependent variable y, from scores of one or more independent variables $x^1, x^2,$ etc. One of the conditions to be met according to Brase and Brase (2017) when incorporating a Pearson product moment correlation coefficient (a numerical measurement which assesses the strength of relationship between x and y) is that there must be a linear relationship between the two variables in question. Cunningham and Aldrich (2012) indicates that data assumptions are the same for both single and multiple linear regression insofar as a linear relationship must exist between dependent and explanatory variables. The regression equation is "the method that minimises the sum of the squared deviations between the regression line and the individual observations" (Conover, 1980, 265). The distinction between the independent and dependent variables is made on the basis that an independent variable cannot be controlled or manipulated, whereas the dependent variable can, notes Bluman (2017). The strength of the relationships between variables increases the predictive power of the model.

The predictive power of the independent variables (pitch, intonation, etc.) in determining the dependent variable (ABRSM grade) is developed through regression analysis. Multiple regression exploits the predictive power of the independent variables to anticipate a linear estimate of the value of the dependent variable, and is a stronger statistical tool than correlation. The gradient of the regression line, from the point of intercept onwards, is plotted between collective residuals for goodness of fit. The residual sum of squares (SS_R) represents the degree of inaccuracy of the regression line. Each predictor variable has its own regression coefficient $b_{1 \text{ to } n.}$ A bi-variate linear regression model was chosen to examine the extent to which variable totals percentage impacted on overall grade results percentage, and scatter plots were developed for each explanatory variable. Hence the regression equations used in this part of the analysis are:

Y = α(constant) + b_1 $X_{1 \text{ (PITCH)}}$;

Y = α(constant) + b_2 $X_{2 \text{(INTONATION)}}$;

Y = α(constant) + b_3 $X_{3 \text{(RHYTHM)}}$;

Y = α(constant) + b_4 $X_{4 \text{(INSTRUMENT)}}$;

Y = α(constant) + b_5 $X_{5 \text{(TRAJECTORY)}}$;

Y = α(constant) + b_6 $X_{6 \text{(DISTANCE)}}$;

Y = α(constant) + b_7 $X_{7 \text{(GUIDANCE)}}$;

Y = α(constant) + b_8 $X_{8 \text{(READING)}}$,

In these equations, alpha is the constant (i.e. the intercept of the regression line) and b_1, b_2, etc. are the regression coefficients for the individual independent. In the instance of this study, Y = the grade achieved, while X_1 to X_8 represents the independent variables, from pitch to reading. At this stage of the analysis, we are looking at the impact of the independent variables in turn, through a series of bi-variate regression equations. In the following section, multi-variate regression analysis will be undertaken, with a regression equation containing all the variables. Hence the multi-variate regression equation will be Y (Grade achieved) = α(constant) + b_1 X_1 (Pitch) + b_2 X_2 (Intonation) + b_3 X_3 (Rhythm) + b_4 X_4 (Instrument) + b_5 X_5 (Trajectory) + b_6 X_6 (Distance) + b_7 X_7 (Guidance) + b_8 X_8 (Reading).

The explanatory and response variable relationships tell us something about the predictive power of the model. In place of certainty, probability and best estimate predictions provide a credible forecasting ability of the likely outcomes. The linear regression model, clusters the data as can be seen in Figure 26 below. From this model, we can infer correlation and predictive power between the two paired data values. Assumptions must be made to determine whether outliers should be excluded before undertaking a regression line.

The first scatter plot shows a positive linear correlation between paired data values of total variable percentage scores and grade percentage score outcomes. The explanatory variable (which in this model it is made up of an amalgam of the observation variables reduced to a single percentage) and the response variable (consisting of grade results made into a percentage) were clustered into the top right area of the chart. A sample correlation coefficient of r=.558 was recorded. From the Cartesian axis we can predict, with significance at the .01% level, how high scores on the predictor variable are likely to reflect similar high scores in grade outcomes. The mean for the accumulated variable percentage scores and grade percentage scores were 81.46 and 78.55 respectively, with standard deviation of 7.837 and 6.933 respectively.

The least squares criterion was used to find the linear equation which best represents the data points in the model. Minimising the squares of the distances between the data points, the regression line gives an indication of the accuracy of the model to predict the response variable. The extent to which the data points of the ordinal pairs of x and y variables are close to the slope line b is an indication of the degree of linear relationship within the model. The fractional part of the variability in y associated with the variability in x is estimated using the least squares criterion, and is represented by the vertical distance d of all data points to the line. The slope b tells us how many units y (the grade) changes for every change in the x units (the independent variables)

The accuracy of the predictive strength of the model is influenced by several components. Marginal changes in the response variable are altered by the inclusion or exclusion of data points on the fringes of the plot. The higher the r value of the correlation coefficient (being closer to 1 or -1), the better the model will fit. The value of the coefficient of determination of r^2 (the square of the sample correlation of r as a measure of proportion of variation in y) is also a consideration. Predictor variables should be interpolated *within* the range rather than extrapolated outside of it, to predict y outcomes. Data should cluster around the least squares line for the model to be a suitable predictor of y outcomes. In the scatter plots which follow, data points representing 37 participant scores are used to generate a regression line between independent variables (the objective scores) and dependent variable (the subjective grade mark).

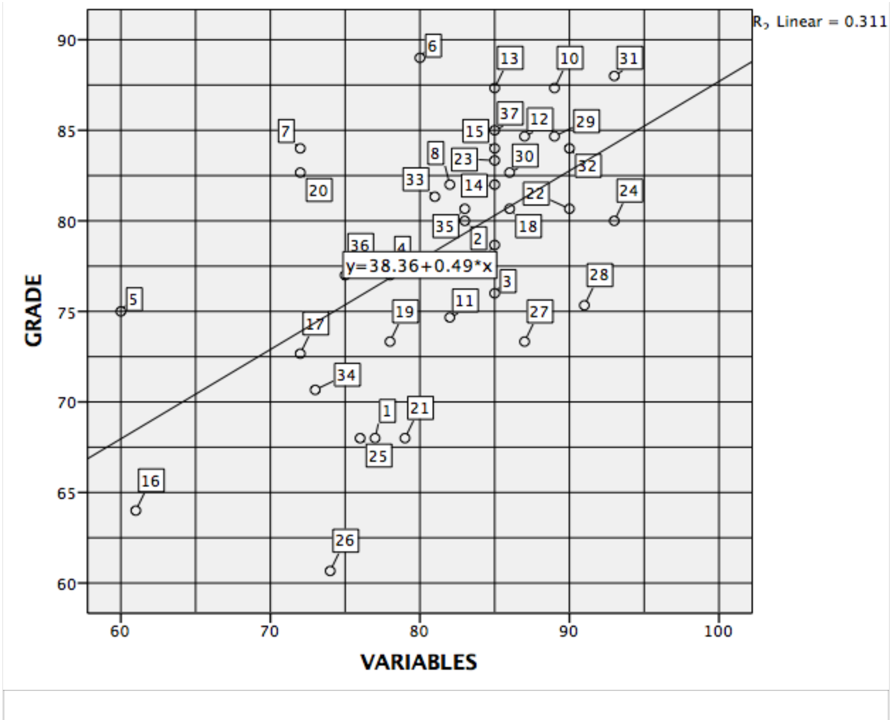


Figure 26. Bi-Variate Linear Regression (Amalgamation of Variable Totals)

In regression, a regression line is created to predict the relationship between two (bivariate) or more (multi-variate) variables and it can be used to identify to what extent variables (X_1, X_2, etc.) can predict the value of the outcome variable y (in this instance, the grade). The prediction equation gives the least squares difference and the refined R_2 coefficient. All variables were clustered in the same area of the plot and the data fulfilled criteria necessary for regression modelling mentioned earlier. The explanatory variable made up of eight percentage explanatory variables collectively yielded a value of $y = 38.36 + .49*x$. The least squares difference was relatively high, reflecting the span between predictive data points and regression line. The $R_2 = 0.311$ was also relatively low. It can be seen from the scatter plot that some of the observed plots (5, 6, 7, 16, 26) might be considered as outliers.

The model fit is a relatively good, considering the data points are clustered between 60% and 95%. The overall prediction suggests that high percentage scores on the accumulative explanatory variable can anticipate, with reasonable certainty, the level likely to be achieved in the grade percentage.

Regression lines and predictive equations for individual explanatory variables were calculated and plotted, and the results are shown in table 00 below, along with the individual pots of the regression lines and the observations. As can be seen, the explanatory power of the regression equations varied considerably - the regression model with the lowest explanatory power was rhythm, for which the R^2 value was only 0.011 – so only just over 1% of the variation in the grade score was accounted for by variation in the rhythm variable. The strongest regression equation was that for Variable 8 (reading), for which R^2 value was 0.370, meaning that 37% of the variation in the grade score was accounted for by variation in the reading variable. The average of the R^2 values for the eight equations was 0.206, i.e. circa one fifth (20%). Given that all the data for variables 1 to 8 was gathered accurately and reliably by the means of data capture technology, which treats all participants from whom it gathers data in exactly the same way, these results cast doubt on the reliability of the ABRSM grade. However, there may be other reasons for the relatively low explanatory power of the bi-variate linear regression equations, with respect to grade scores. For example, it could be that each of the skills that were measured (pitch, intonation, etc.) are insufficient in their own right to affect grade scores, but that the do have an impact when they act in unison, which is what shall be examined through the causal path modelling.

Variable	Regression Equation	R^2
Pitch	$Y = 69.6 + 0.014*X_1$	0.114
Intonation	$Y = 42.42 + 0.4*X_2$	0.121
Rhythm	$Y = 75.14 + 0.04*X_3$	0.011
Instrument	$Y = 58.7 + 0.26*X_4$	0.358
Trajectory	$Y = 35.59 + 0.51*X_5$	0.237
Distance	$Y = 61.8 + 0.23*X_6$	0.146
Guidance	$Y = 39.85 + 0.41*X_7$	0.288
Reading	$Y = 22.34 + 0.06*X_8$	0.370

Figure 27. Summary statistics for least squared regression

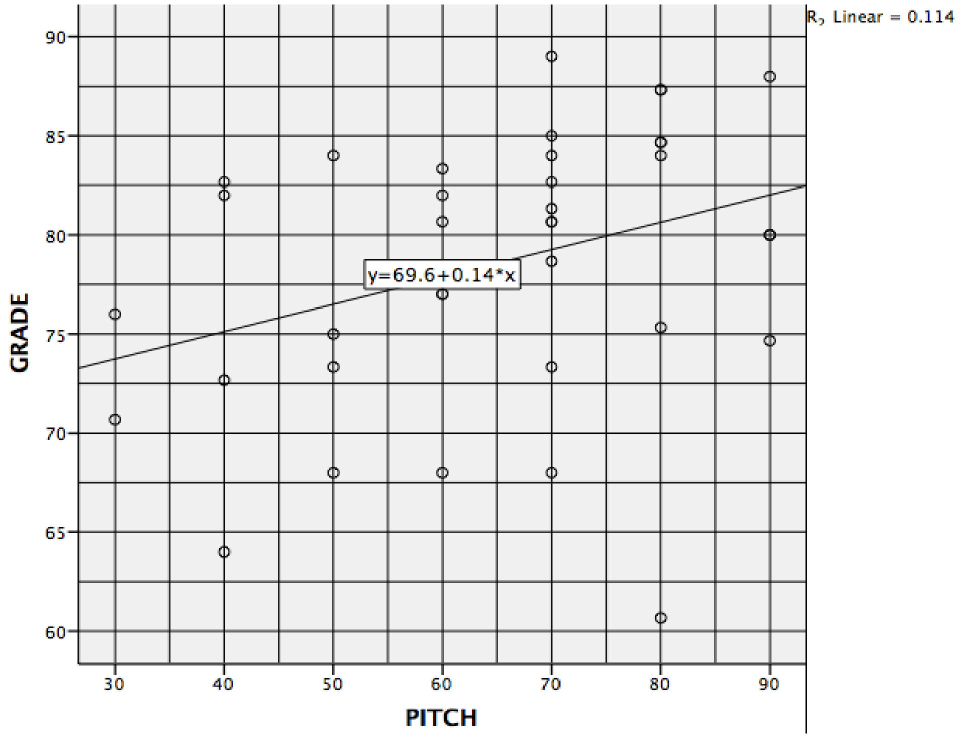


Figure 28. Bi-Variate Linear Regression – Pitch

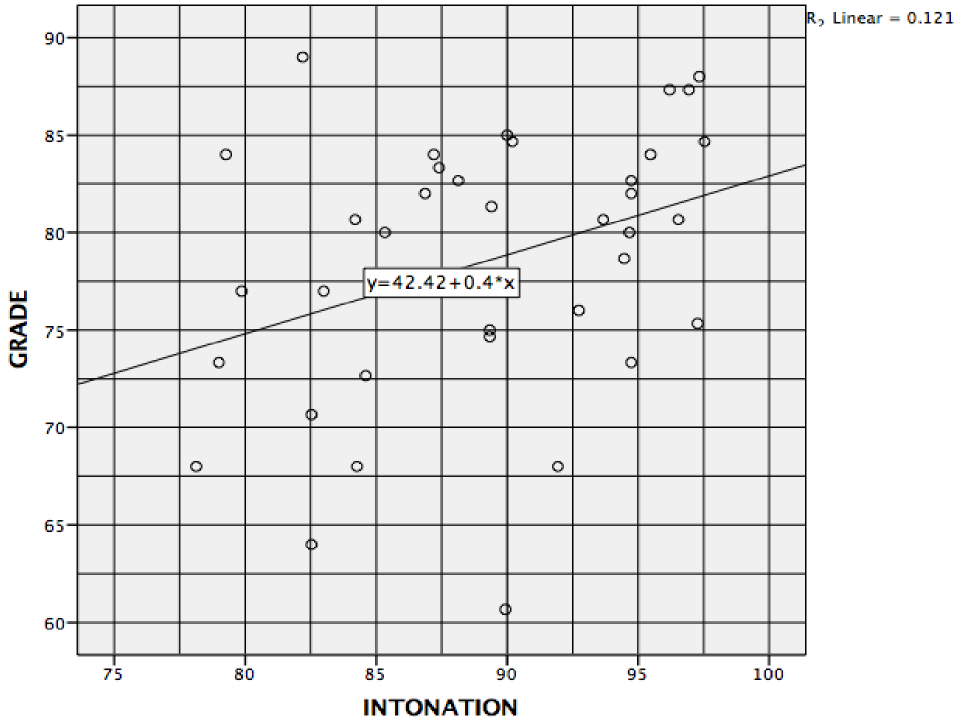


Figure 29. Bi-Variate Linear Regression - Intonation

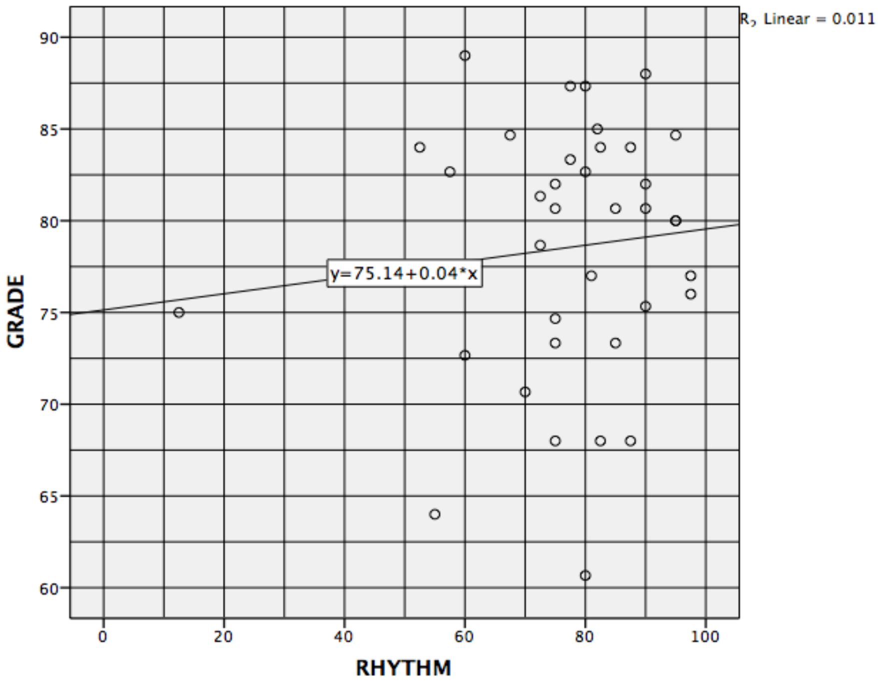


Figure 30. Bi-Variate Linear Regression - Rhythm

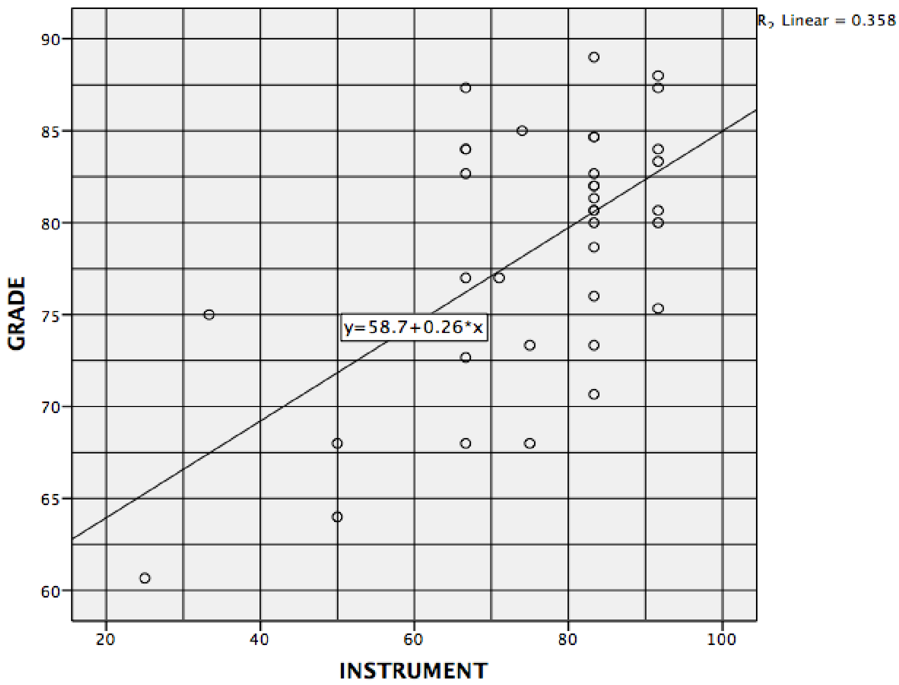


Figure 31. Bi-Variate Linear Regression – Instrument

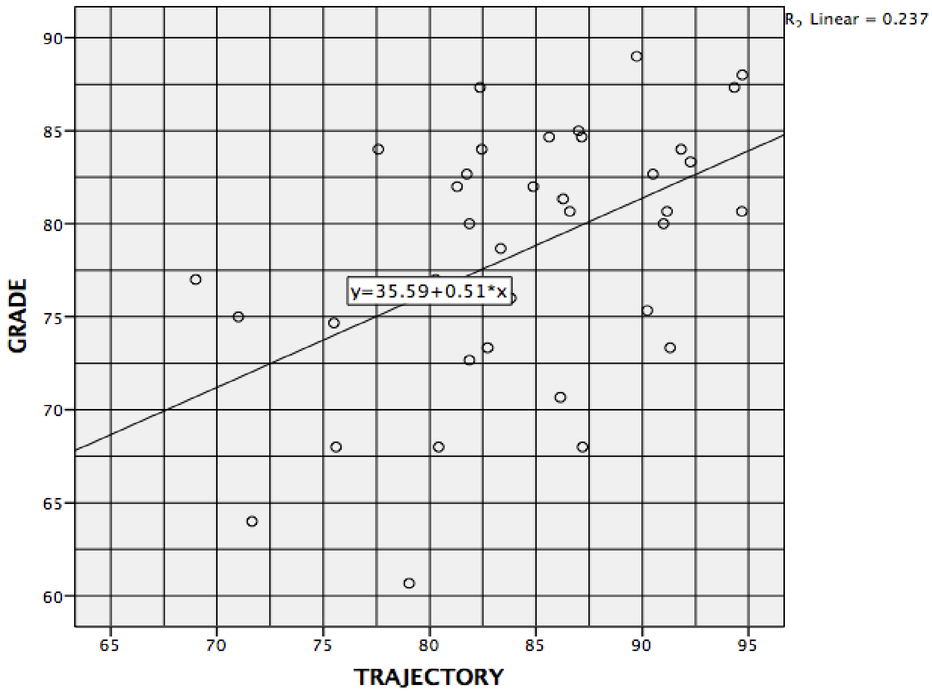


Figure 32. Bi-Variate Linear Regression – Bow Trajectory

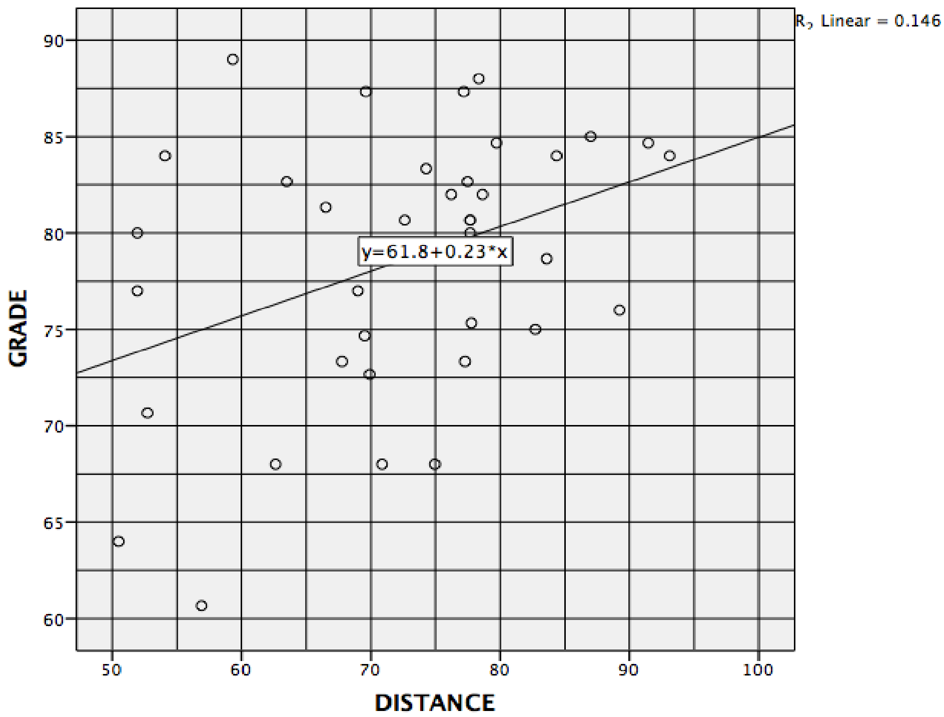


Figure 33. Bi-Variate Linear Regression – Bow Distance

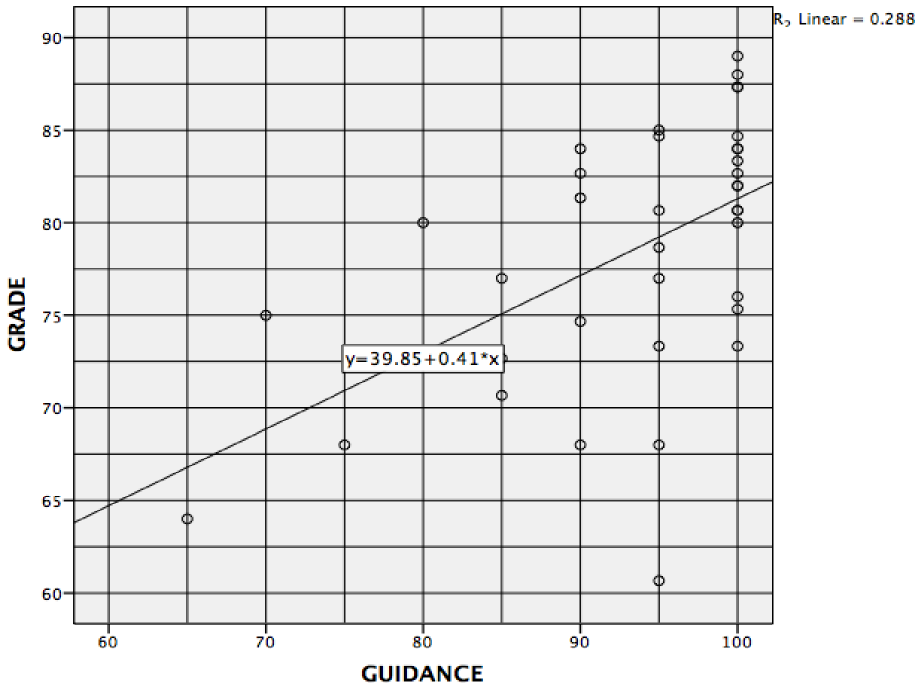


Figure 34. Bi-Variate Linear Regression – Bow Guidance

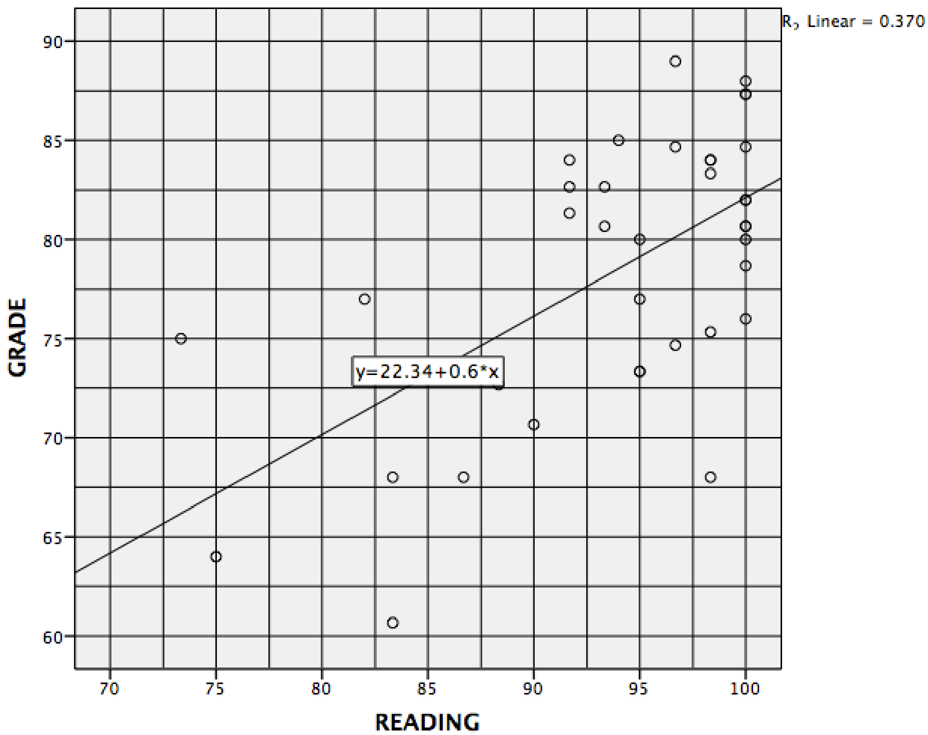


Figure 35. Bi-Variate Linear Regression – Reading

These plots provide a visual appraisal of the bivariate linear regression models for the numerical data presented. Care should be taken when making predictions on the basis of our sample of 37 participants, as the equivalent relationships, when calculated for the population, may have weaker or stronger explanatory power (R^2 values) than was evidenced in the sample. As can be seen from the plots, the data points for all of the eight variables fall largely in the same quadrant area of the graphs, although there is considerable variation in the spread of the plots.

4.6 Building Causal Paths with Multiple Regression

The analysis so far has examined the links between the dependent variable (the grade score) and individual independent variables by means of correlation and bi-variate regression. However, it is evident that the independent variables (pitch, intonation, rhythm, reading, etc.) mesh together and thereby affect the grade score collectively and in unison, rather than individually. For example, without the ability to read music easily, the ability to identify and adopt the right rhythm for a piece of music becomes problematic, Hence, to understand the process of grade determination more clearly and accurately, multi-variate, rather than bi-variate regression, is needed.

In addition to multiple regression analysis, and stemming from the idea of causal modelling, path analysis simplifies the complexity of the underlying latent factors within a multi-variate regressional model. While it gives no indication of causality, path analysis can rule out null causal paths, by using estimates of the significance of the impact of the individual variables, thereby enabling the emergence of a causal model, in which all the independent variables have a significant impact on the dependent variable (that is, the grade). Inter-correlations exist between independent variables (pitch, intonations, etc.) as previous work evidenced in Figure 25 on page 133 shows. However, path analysis tests the strength of relationships amongst the independent variables, and between the independent variables and the dependent ABRSM Grade variable. Using multiple regression equations, we can test the effect of all the independent variables, acting together, on the dependent grade variable, This is done by comparing possible regressional models, to see which one best fits the data.

Causal path models are used in the social sciences to try to understand how, and with what impact, different independent variables can affect dependent variables. The process of creating causal path models can be readily explained by means of a worked example using general educational attainment, as measured by the ability to read among primary school children. It can be hypothesised that the ability to read may be affected by (inter alia) age, gender, IQ, number of siblings, and the reading abilities of mother and father. This hypothesis would lead to the creation of the following model, which is shown below, and for which the multi-variate regression equation will be Hence the multi-variate regression equation will be Y (Reading Ability) = α(constant) + b_1 X_1 (Age) + b_2 X_2 (Gender) + b_3 X_3 (Siblings) + b_4 X_4 (IQ) + b_5 X_5 (Mothers Reading Ability) + b_6 X_6 (Fathers Reading Ability).

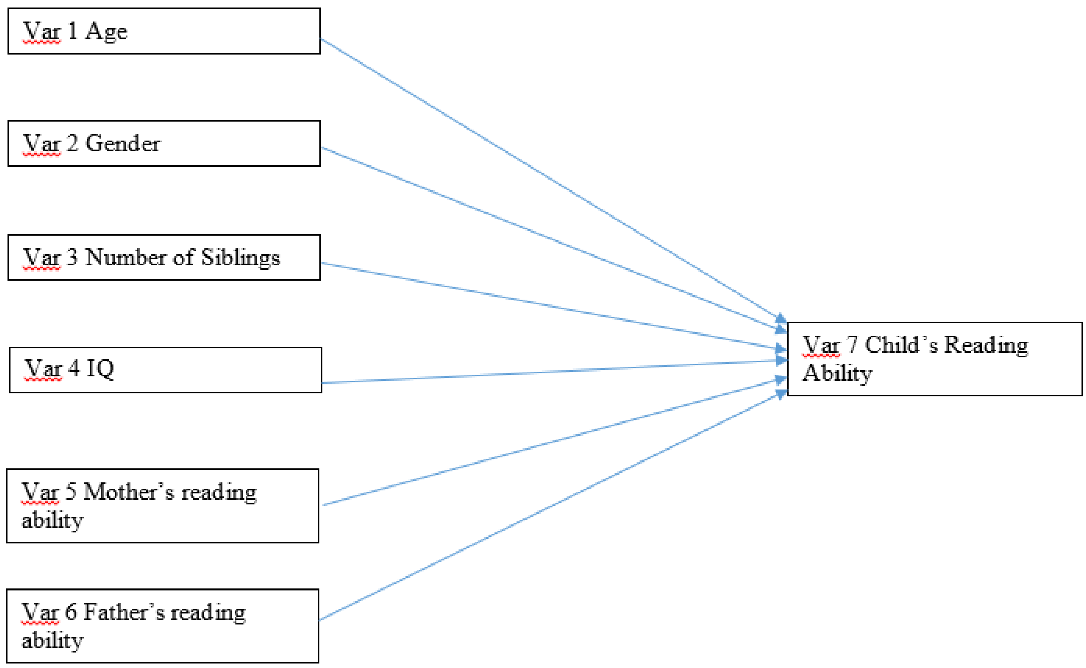


Figure 36. Causal Path Model Example

Using the SPSS Regression procedure, all the variables are entered into the regression equation, to see the impact of the six independent variables on the dependent variable. This initial regression could produce a causal model similar to that shown below, and the regression equation would be: Y (Reading Ability) = α(constant) + 0.38 x X_1 (Age) + 0.11 x X_2 (Gender) + 0.2 x X_3 (Siblings) + 0.61 x X_4 (IQ) + 0.52 x X_5 (Mother's Reading Ability) + 0.20 x X_6 (Father's Reading Ability).

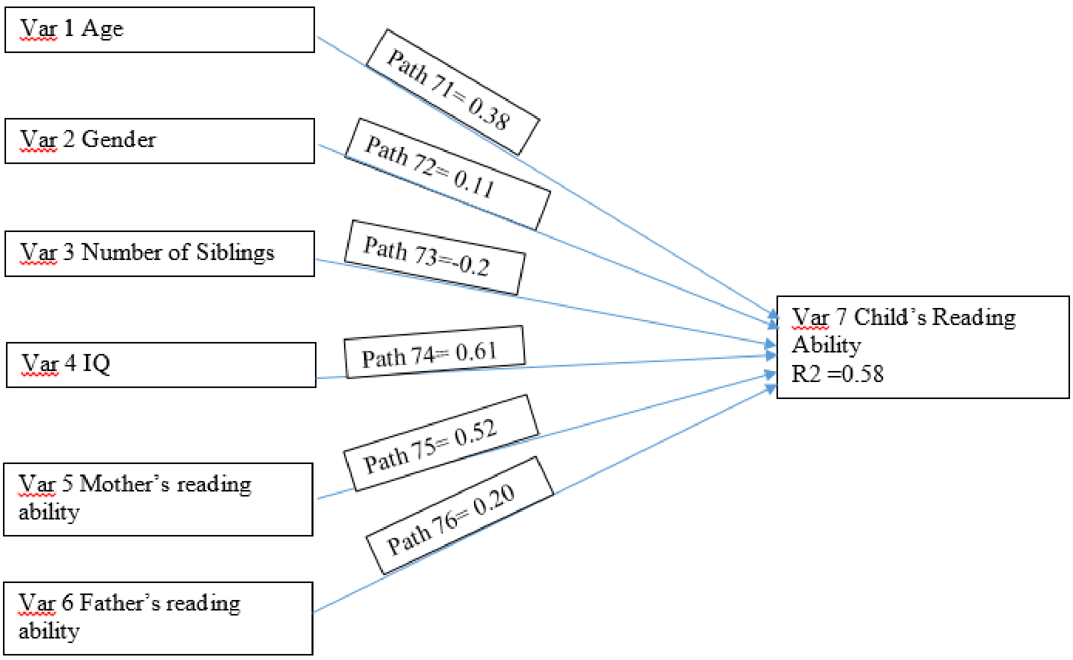


Figure 37. Reduction Sequence Example – Initial Path Values

The R^2 for the model is 0.58, hence acting together, the variables in the model explain 58% of the variation in a child's reading ability. However, the model shows that some of the variables (Gender, Number of Siblings and Father's Reading Ability) each have a relatively small impact, as discerned by the size of the beta weights (that is the standardised regression coefficients). Consequently, these variables are removed from the regression equation, and it is then re-calculated with the remaining variable and generates the following multivariate regression equation Y (Reading Ability) = α (constant) + 0.43 x X_1 (Age) + 0.54 x X_4 (IQ) + 0.55 x X_5 (Mother's Reading Ability) causal path model. As can be seen in Figure 38 below, the overall explanatory power of the model has dropped slightly, from .58 to .47. However, we have clarified the model and shown that Age, IQ and Mother's Reading Ability are the most important factors.

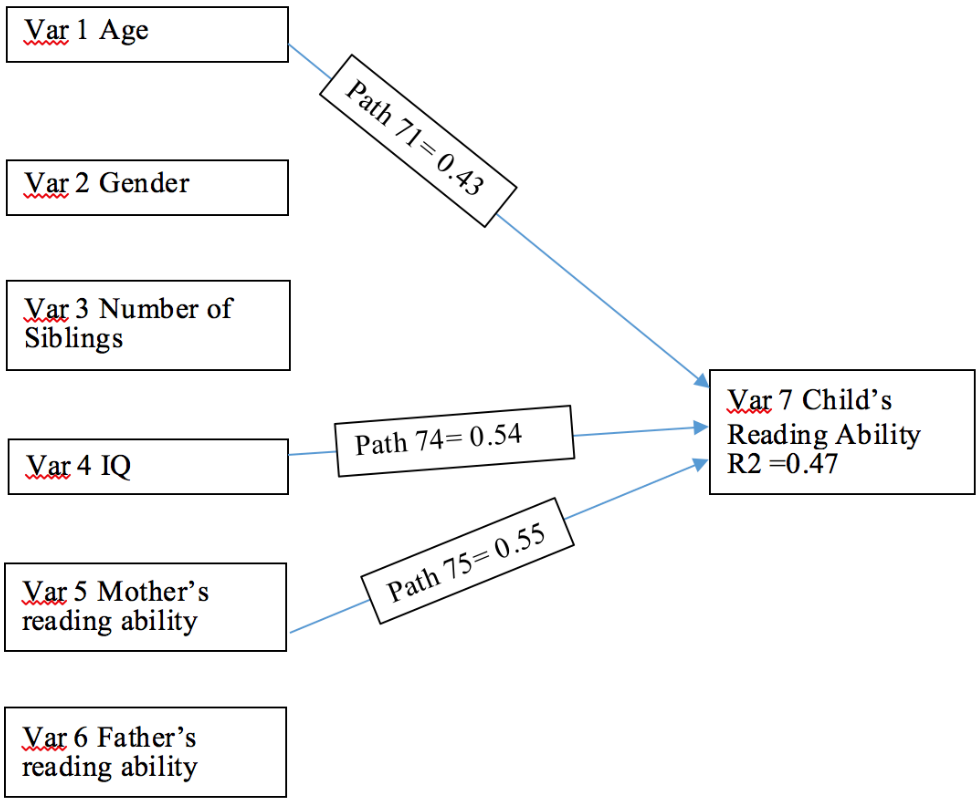


Figure 38. Final Causal Path Model Example

As can be seen from this example, the goal of extrapolating which dependent variables predict an independent variable best, as in multiple regression, is taken a stage further with path analysis. Causal path analysis, and the multi-variate regression equations that it utilises, sheds some light on possible cause and effect relationships, and is concerned with establishing the presence of patterns in the data. Streiner (2005), aligning the term 'exogenous' with independent and 'endogenous' with dependent variables, suggests that exogenous variables happen outside the model, while endogenous ones happen within it. He states that analysis is achieved by estimating "1) the paths, 2) then covariance among the exogenous variables, and 3) the variance of the exogenous variables but not the variances of the endogenous ones." (p. 120). He also states that the technique is concerned with testing an existing model rather than building a new one.

Altering the model to understand which variables best predict an outcome is achieved by understanding what affects exogenous variables have, and the strength with which they correlate with one another, so as to eliminate each inconsequential variable in a variation of the model. Hence the multi-variate causal path analysis which will be used to examine the determinants of ABRSM grades will utilise the following regression equation in which Y (Grade achieved) = α(constant) + $b_1 X_{1(\text{PITCH})} + b_2 X_{2(\text{INTONATION})} + b_3 X_{3(\text{RHYTHM})} + b_4 X_{4(\text{INSTRUMENT})} + b_5 X_{5(\text{TRAJECTORY})} + b_6 X_{6(\text{DISTANCE})} + b_7 X_{7(\text{GUIDANCE})} + b_8 X_{8(\text{READING})}$.

4.7 Causal Path Modelling for Grade Determination in Violin Playing

In the educational setting, causal path modelling as described by Karadag (2012) indicates a technique which helps to explain the extent to which certain variables influence a dependent variable more than others. Interrelated variables, relating to posture and bow angles have been reduced to a single trajectory variable, as they measure aspects of the same underlying dimension. Causal path analysis described by Streiner, (2005, 118) shows how graphical indications of the way variables fit together help to explain variation in the strength of the impact of the independent variables on the dependent variable. Path analysis involves changing the models to see which one best fits the data, in terms of (for example) maximising the R^2 value or minimising the number of independent variables. The models chosen for this study included the intuitive (based on the researcher's skills and knowledge of the process of teaching the violin), the inclusive (which maximises the number of independent variables) and the parsimonious (which maximises the R^2 value of the model), to analyse latent exogenous variables and the strength of their impacts on the endogenous variable. Rejection of the $H_0 : \beta=0$ "implies that at least one of the regressors … contributes significantly to the model" (Montgomery, 1992, 135).

The intuitive model entered the independent variables into the regression equation in an order deemed to be of most significance based on extensive teaching experience, string learning theory, and a reading of the literature, using the enter option in the SPSS Regression procedure. The order of entry reflects the importance placed on individual independent variables, such as the need for a correct bow trajectory, or a literacy framework on which the teaching of music can be based. Consensus is widespread regarding core principles within an instrument discipline and several decades of experience, teaching both privately and in schools, bear testimony to this. Core understandings, such as the need to cultivate a correct posture during the learning process and the development of a precise intonation, make up two cornerstones of such a consensus. Such assumptions, discussed at length in earlier chapters, frame the backdrop to the intuitive model. In designing these models, 'specification error' – described by Berry (1985), where the functional form of the relationship is stated improperly or where the wrong independent variable is used – which necessitated careful planning to guard against the potential for these errors to take place.

The inclusive model has the highest explanatory power. The stepping method criteria - probability of F entry level was set at 0.05 and the removal criteria was set at 0.1. The Forward inclusion option in the SPSS Regression procedure causes variables to be added to the equation individually, and at each stage the smallest probability-of-F value is assessed with respect to the entry criterion. In this way, the improvement of the R^2 value of the equation is assessed in terms of the impact of adding each successive variable. This model, while having higher explanatory power, also had many independent variables. The inclusive model maximises R^2 by including each variable one by one, if they meet the pre-set threshold value.

Parsimonious models were developed to deduce bow trajectory elements, reading elements and all variables combined to discover the attributes which were having the largest effect. The variables are deleted progressively when shown to have a very limited effect, using the Backward Exclusion option in the SPSS Regression procedure. Using this procedure, all variables are included initially, and each is assessed in terms of its contribution to the model. The variable with the lowest contribution is removed, and the equation is re-run. This minimises the number of mistakes which are likely to occur in the model until all remaining variables have a significant *p*-value – that is the all make a significant independent contribution to overall explanatory power (R^2). In these models, the R^2 value may not be as high as in other models, but with fewer independent variables included, it becomes more apparent which variables are likely to be having a greater impact, as the variable with the largest probability-of-F value is removed when the value is larger than the removal criterion set at .10, which is standard in SPSS.

Bryman (2001, 248) states that path analysis "cannot be used as a substitute for the researcher's views about the likely causal linkages among groups of variables".

Hence, a strong theoretical premise anticipated that Pitch Discrimination (Variable 1) would have an effect on Intonation (Variable 2). Similarly, it was anticipated that Rhythm (Variable 3) would have a significant effect on Reading (Variable 8). To a lesser extent Instrument Angle and Guidance were also anticipated to have a direct impact on the Grade result. The stronger causal paths in the regression model are shown in red.

4.7.1 The Intuitive Model

In the Intuitive model, the explanatory power for the impact of the independent variables, Pitch, Distance and Trajectory, on the dependent variable Intonation have a value of $R^2 = .404$. The explanatory power for the independent variables Intonation, Reading, Instrument and Guidance on the dependent variable (Grade) was $R^2 = .427$. The R^2 values are indicated on the subsequent Inclusive and Parsimonious models.

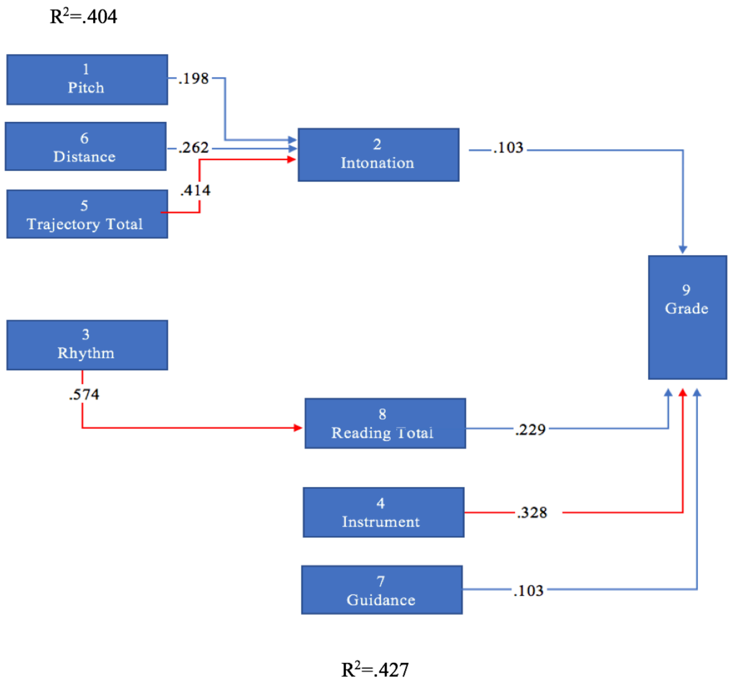


Figure 39. Intuitive Causal Path Model

The intuitive model, while being less parsimonious than other models, nevertheless tries to predict implicit variables, through ones which are observable, by displaying "whether the pre-determined relation pattern can be proved or not by the obtained data" (Karadag, 2012, 196). The direction of the predictor or exogenous variables' effects on other exogenous variables is not determinable within the model, although an inference supported by theory can make a strong case for any interpretation of the data within the model. The causality, however, is not of paramount concern, as the regression coefficients, which support the alternative hypothesis, satisfy the aspirations of the research question, which postulated a link between objective and subjective methods of assessment.

In this model, a clear association is made between observable heterogeneous qualities which underlie factors known to impact positively on performance (ABRSM grade). The way in which the trajectory variable affects the intonation variable warrants further discussion in this regard. It reiterates a discovery made at the pilot stage of the research about counter-intuitive associations, such as the one between bow distance and intonation, and has given rise to a theory which helps to explain these associations on page 72.

Structural equation modelling SEM according to Karadag (2012) is dependent on theory to rationalise the way in which predictor variables are organised in the model. Theory is also needed when drawing conclusions about how the model fits the data. Theory being relied on in the intuitive model draws on theoretical assumptions about the relationship between pitch determination and intonation alluded to in Gerle (1983). Also, an association between posture conventions and tone production are brought to the fore. Sight-reading ability is understood to be synonymous with good musicianship according to Kopiez and Lee (2008), and it enables the practitioner to engage with the literature on different levels notes Green (2002), accessing established works and repertoire known to have educational, technical, and pedagogical merit, creating new possibilities for educational Bildung in music education to function.

4.7.2 The Inclusive Model

The Inclusive Model is derived by putting all the variables into the regression equation to start with, i.e. Y (Grade achieved) = α(constant) + $b_1 X_1$ (pitch) + $b_2 X_2$ (Intonation) + $b_3 X_3$ (Rhythm) + $b_4 X_4$ (Instrument) + $b_5 X_5$ (Trajectory) + $b_6 X_6$ (Distance) + $b_7 X_7$ (Guidance) + $b_8 X_8$ (Reading). Then the contribution of each independent variable (pitch, intonation etc.) is assessed. The independent variable with the smallest contribution to the determination of the dependent variable is identified, as this will have the smallest *b* regression coefficient, this variable is then removed, and the regression equation is re-run. This process continues until all the remaining independent variables left in the regression equation have a statistically significant impact on the dependent variable. The results of this process are shown in Figure 40 below. Moving from left to right, each vertical column represents a new model in which one variable has been eliminated. When the elimination takes place, the remaining variables in the model register an altered beta value when the regression equation is re-calculated, in accordance with the elimination criteria stated. The conclusion of this process establishes the most parsimonious elements present from the all-inclusive variables array.

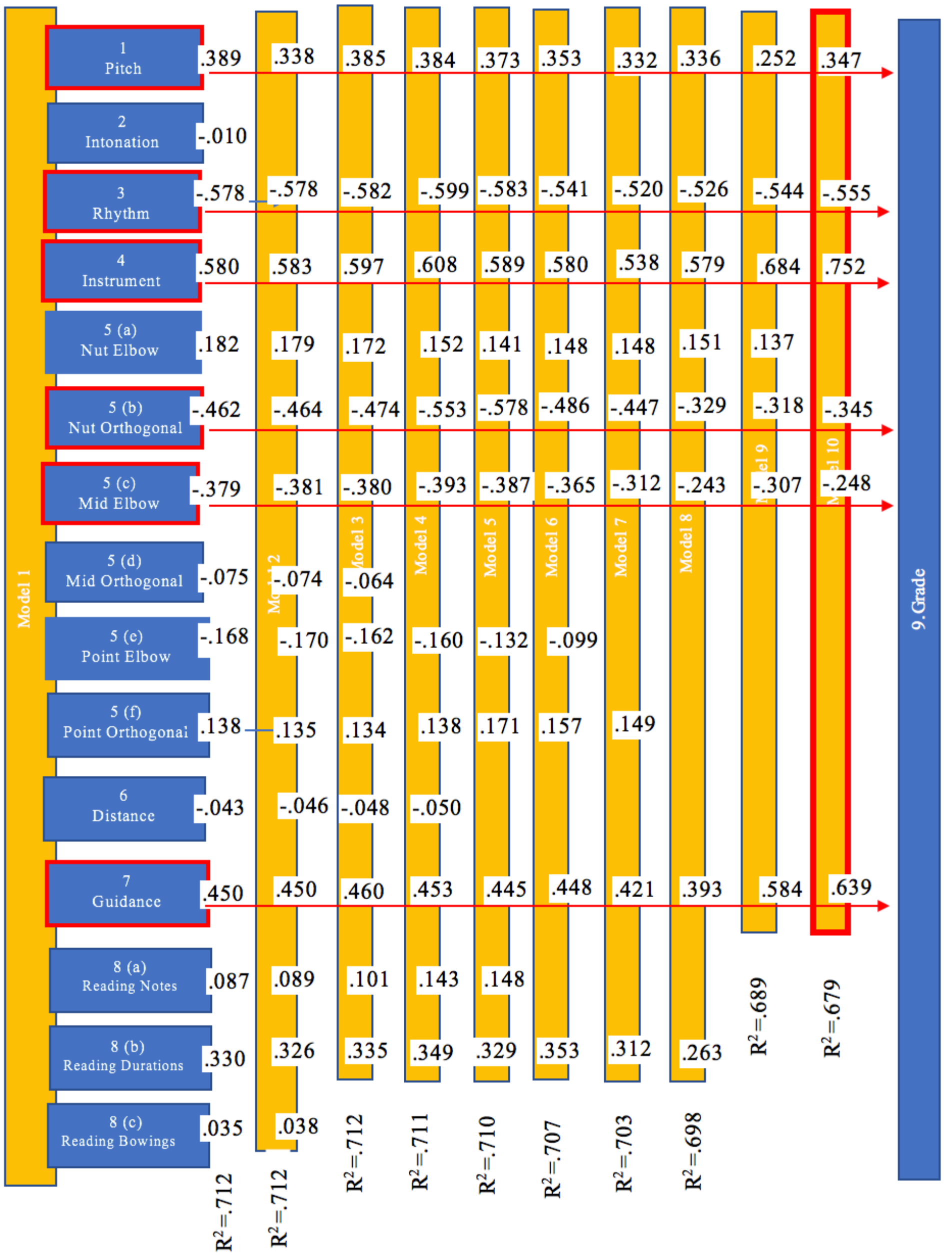


Figure 40. Inclusive Causal Path Model

Hence, the Inclusive model was created using backward exclusion, with all variable equations on the left producing an R^2 of .712. Then the intonation variable was removed as it had the smallest beta weight (.010), and the equation was re-run. Finally, a regression equation emerged in which six independent variables (Pitch, Rhythm, Instrument, Nut Orthogonal, Mid Elbow, and Guidance) had a statistically significant impact on ABRSM grade, and which produced an R^2 of .679. Hence this multivariate regression equation explained 67% of the variation in the ABRSM grades.

Postural or physical qualities, which are embedded in the six trajectory/bow observations in Variable 5, are looking at something which is observed independently of any audio analysis, such as that made in Variable 2, relating to intonation accuracy. These tasks are undertaken from separate cognitive and cerebral domains without any logical association. Over many years teaching novice string players, however, it was noted how many students with relatively good bowing technique, reacted to incorrect intonation by constricting bow travel distance or altering the bow trajectory, to compensate or diminish the effect of the objectionable sound, as a reflex.

More importantly, as a string teaching practitioner, it was noted how many players who displayed this tendency seemed to have difficulty correcting it, after the intonation problem had been addressed. In short, 'bad' intonation seemed to exacerbate bow trajectory, yet corrected intonation did not automatically reverse this. Intonation defects shown to correlate in this way with bow trajectory suggest epistasis based on a theory of observed association. It is more likely to be true, however, that the association works in both directions, given that when someone is confident and secure with their intonation, they should also feel more committed to driving the bow across the strings effectively to express this confidence, creating the opposite cause and effect. Further research in this area could help clarify the extent to which these associations can be supported.

Analysis has shown that variables which have the strongest effect on grade outcomes, according to the size of the regression coefficients to be: reading, instrument, distance, trajectory and guidance, as all the β weights for these variables are significant at the 5% level. Variables that failed to meet this inclusion threshold included intonation, pitch and rhythm.

4.7.3 The Parsimonious Models

Analysis using the 'Backward Regression' procedure in SPSS was conducted to develop 3 parsimonious models: (a), (b) and (c). Parsimonious Model (a), related to Bow Trajectory sub-variable components a-f, and endogenous variable Grade Outcome. Model (b) related to reading components and Model (c) related to variables which remained at the end of the inclusive model elimination process mentioned earlier. The Backward removal method puts all the independent variables into the equation and removes predictor variables with the highest p value (lowest significance) one at a time, altering the standardised beta coefficients in the process.

Parsimonious model (a) (shown in Figure 41) extracted the strongest sub-variable bowing components in Variable 5. Variable 5 is constructed from 6 sub-variables relating to bow trajectory shown below. The value of $R^2 = .44$ for the six elements in this variable could be explained by two of the components: Nut-Elbow and Point-Elbow which alone produced an R^2 value of 0.392.

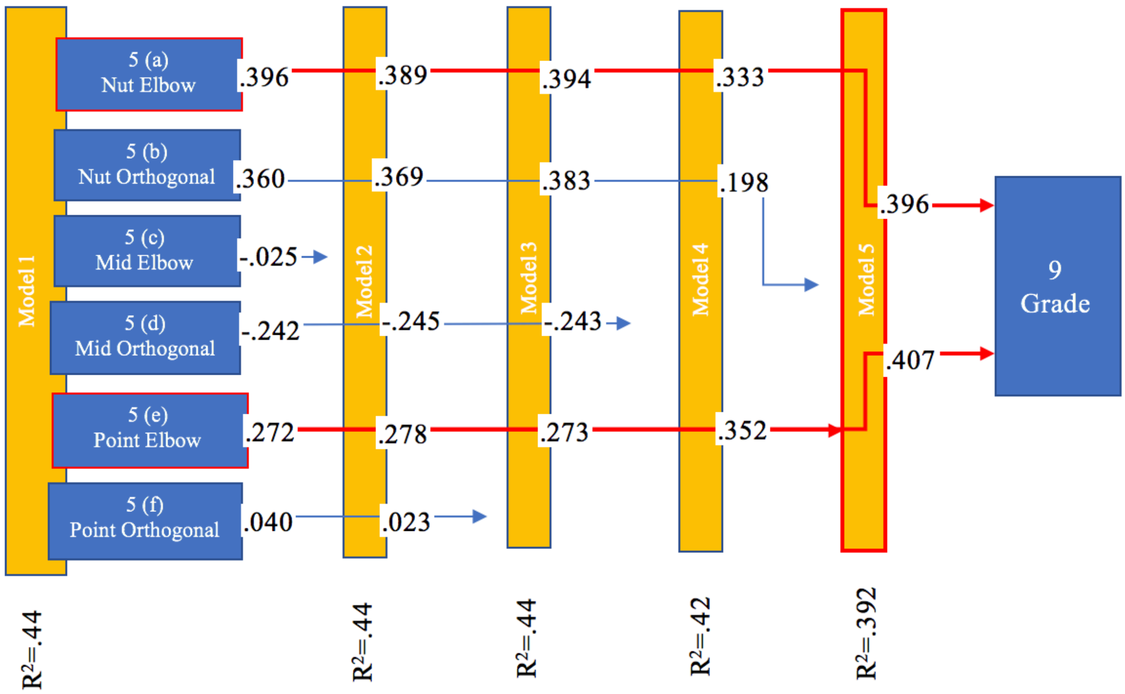


Figure 41. Regression Results – Parsimonious Model (a)

Parsimonious Model (b) (shown in Figure 42 below) establishes parsimony amongst the three sub-variables relating to reading abilities: Note Pitches, Note Durations and Bowing Indications. Again, using the Backward Elimination criteria of least significant variable, the Note Duration variable produced the largest beta coefficient (.654) recorded.

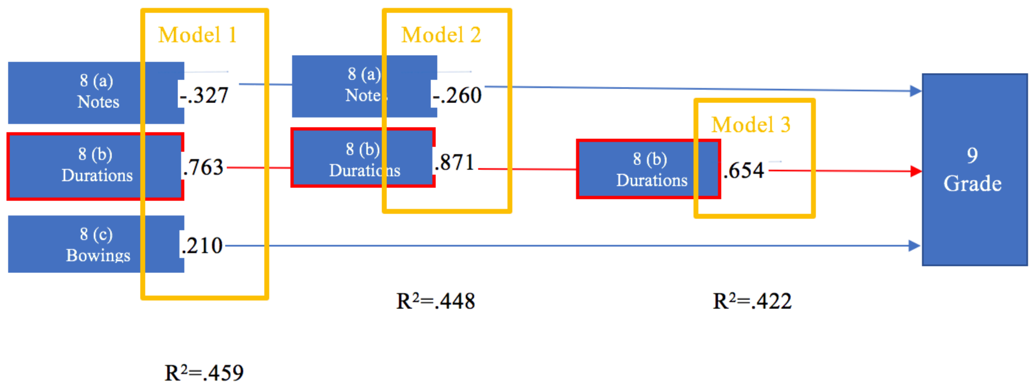


Figure 42. Parsimonious Model (b)

The Parsimonious Model (c) (shown in Figure 43) illustrates the evolution of the most parsimonious model also through backward elimination. In this model, the variables not eliminated in the inclusive causal path model in Figure 40 on page 152 are re-run between the grade outcome variable. The negative beta values appear to reflect negative score input data rather than percentage score calculations made for comparison and correlation.

	Model 1	Model 2	Model 3	Model 4
Nut Elbow	.151	.151	.137	
Point Elbow	-.093			
Reading Durations	.229	.263		
Pitch Detection	.356	.336	.352	.347
Rhythm Accuracy	-.546	-.526	-.544	-.555
Instrument Angle	.620	.579	.684	.752
Nut Orthogonal Angle	-.359	-.329	-.318	-.345
Mid Elbow Angle	-.289	-.243	-.307	-.284
Bow Guidance	.417	.393	.584	.639
Explanatory Power	R^2 =.702	R^2 =.698	R^2 =.689	R^2 =.679

Figure 43. Parsimonious Model (c)

Fisher (1925) advises that the elimination of variables in causal path models requires a qualitative interpretation. This is reiterated by Wright (1934), stating that the purpose of path coefficients is to “combine the quantitative information given by the correlations with such a qualitative information as may be at hand on causal relations to give a quantitative interpretation” (p. 193). While epistasis showing the dominance of one variable over another is not present in the data, it can be said that all of the factors affect in some way the grade outcome to varying degrees. The emphasis placed on variables in relation to strategic pedagogic objectives may be a possible area for future analysis research. The myriad of cognitive and cerebral considerations involved in the learning of a musical instrument necessitate multiple factors to be at play in the learning at any one time. The findings of this parsimonious model do not fully concur with intuitive assumptions. The significance of this model, therefore, is inconclusive as a qualitative interpretation which rationalises the inferential statistic in this instance cannot be made.

4.8 Feedback comparison

In Figure 44 on page 157 all of the data points are assigned to each participant numbered at the bottom of the page starting with participant 38. We can see how the individual instrument specific areas of learning can be compared for each individual. We can also see the degree to which individual variables fall below the grade result shown here in in blue. These sub-blue items are of interest to the teacher as evidence shows how higher levels have a positive bearing on the grade results. Linear regression analysis described on page 135 illustrates the predictive ability of individual variable scores to anticipate grade outcome. By raising these low scores during learning, the analysis suggests that this will increase grade result.

4.9 Graphical Representation of Individual Elements

Figure 43 below illustrates the comparison between predictive independent variables pitch, intonation, rhythm, instrument, angles, distance, guidance and reading, and the dependent variable: the ABRSM grade result. When viewed together, the variable scores and their relationship to grade result appears random and inconsistent. By separating out the individual variables and placing them alongside the endogenous grade result, a clearer picture of the link between the objective component and the subjective grade response becomes more apparent.

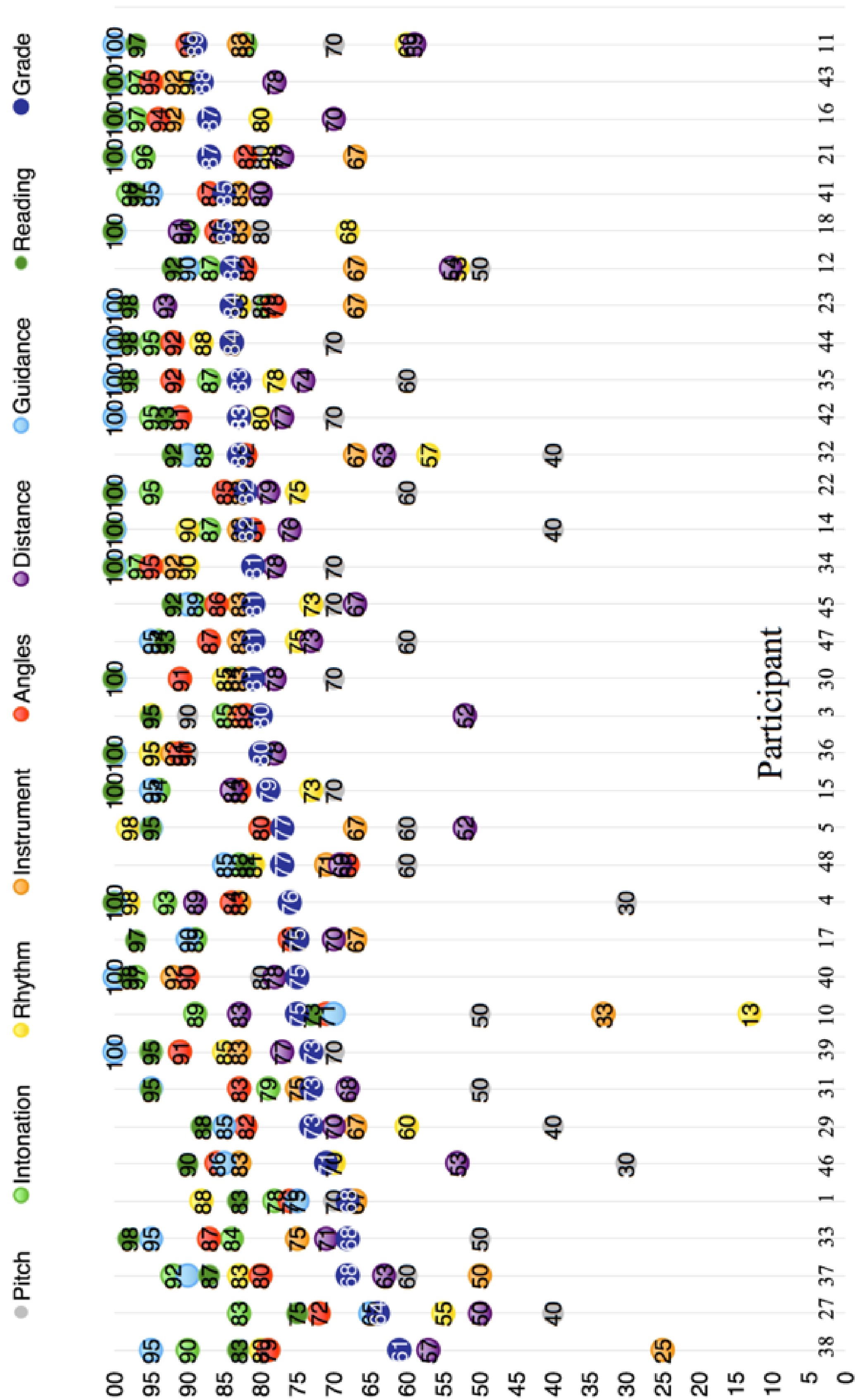


Figure 44. Graphic Representation of All Variable Data and Grade

For this illustration, the participants are presented along the x axis and the percentage scores for exogenous variables and endogenous variable are presented along the y axis. In this representation, the sample data is contextually grounded in a tangible comparative display which can be utilised directly to isolate and make determinations about a participant's progress based on their strengths and weaknesses.

To best understand these graphs, it should be stated at the outset that the observation constructs are grounded in string teaching theory, which supports the constructs mentioned in the methodology chapter. Assumptions which stem from many years of teaching within a Western art music tradition concur with these associations. For instance, one should notice if something is out of tune (pitch perception), before one might be expected to play in tune (intonation). The graphs bring to the fore an interpretation of the data, which is tangible and of direct benefit to student and teacher.

The data percentage scores in the y axis are tied to individual participants along the x axis. This makes explicit a relational context where deficits can be found. For instance, where the grey trace line relating to pitch performance descends below the blue line relating to grade result, we have an indication where student and teacher must focus in relation to that learning objective. As was apparent with the correlation, regression prediction lines, and path analysis models discussed earlier, a positive linear relationship exists between the explanatory and response variables. The degree to which a subsequent explanatory variable score change alters a response variable change was explored through linear regression modelling. Visual representations can be acted on directly in the field and provide a valuable resource for teacher and student. Looking at Figure 44 above representing all participants, it is possible to determine strength and weakness for each participant in relation to each variable.

At a glance, the practitioner can build on this format of data presentation to address the various aspects and deficits in the learning process needing attention. In addition, the order in which interventions should take place is made clearer. For instance, Participant 32 has a low score in pitch detection and poor ability to differentiate between 'in' and 'out of' tune. It is logical that pitch detection exercises should be done before improvements with intonation can be expected. The correlation between pitch detection and intonation is moderate, with correlation of r=.381 (significant at the 5% level) but string teaching theory such as Fleisch (1939) supports the assertion that, in practice, they are inextricably linked. Similarly, with Participant 3, it can be expected that improvements to bow trajectory angles will improve learning outcomes for this student.

The grade percentage outcomes and the variables mean percentage outcomes when viewed together mirror an erratic link between the two approaches to assessment. The amalgamated variables (in red) show varying degrees of excellence achieved by each participant. When viewed alongside incrementally increasing grade results for those students achieved through subjectively generic accounts (in blue), it becomes apparent how areas of difference between the two plots highlight important areas where real learning potential can take place in the spirit of educational Bildung. Places where the red line is below the blue line suggest the most likely area where grade improvement potential may exist, while areas above suggest strengths which may be less understood for their pedagogic value in 'musical outcome' analysis and subsequently not reflected very accurately in a grade outcome mark.

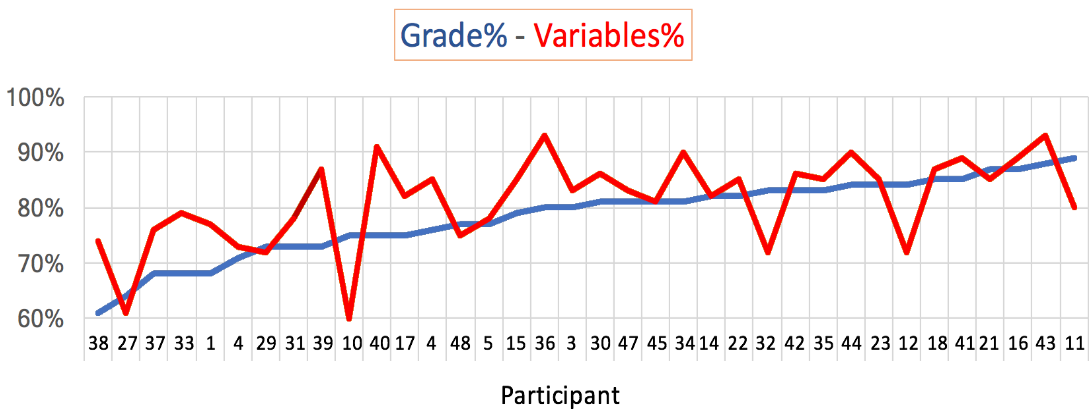


Figure 45. Grade% and Accumulated Variables% Comparison

The order in which different aspects of novice string learning can be addressed is crucial, given the cognitive overload concern in relation to student ability and temperament. While each aspect will be addressed in a different order depending on teaching style and student progress, a chronology which minimises this problem is put forward.

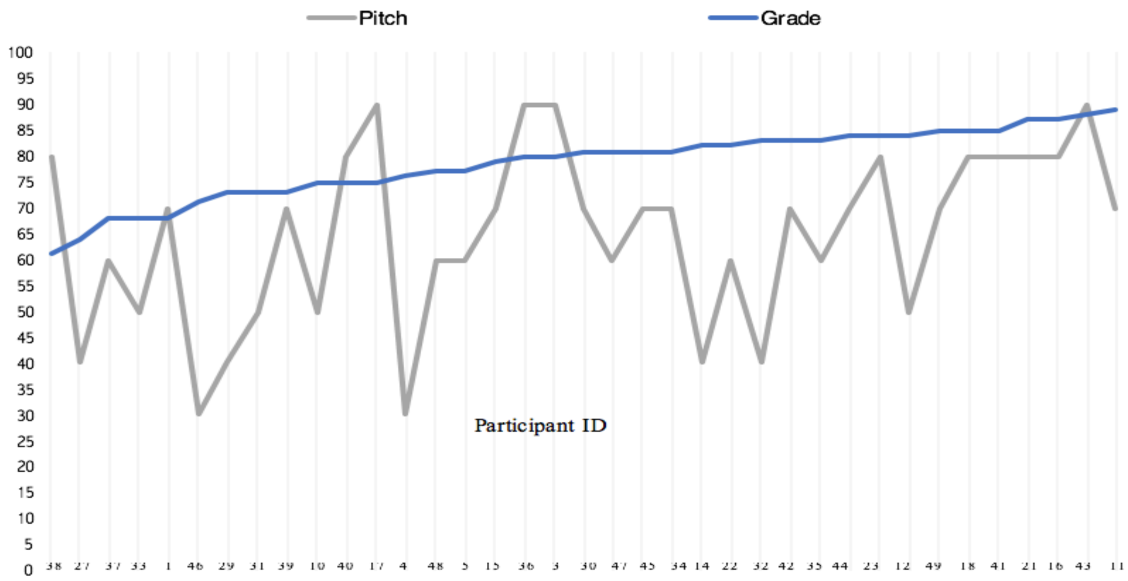


Figure 46. Graphic Representation – Pitch

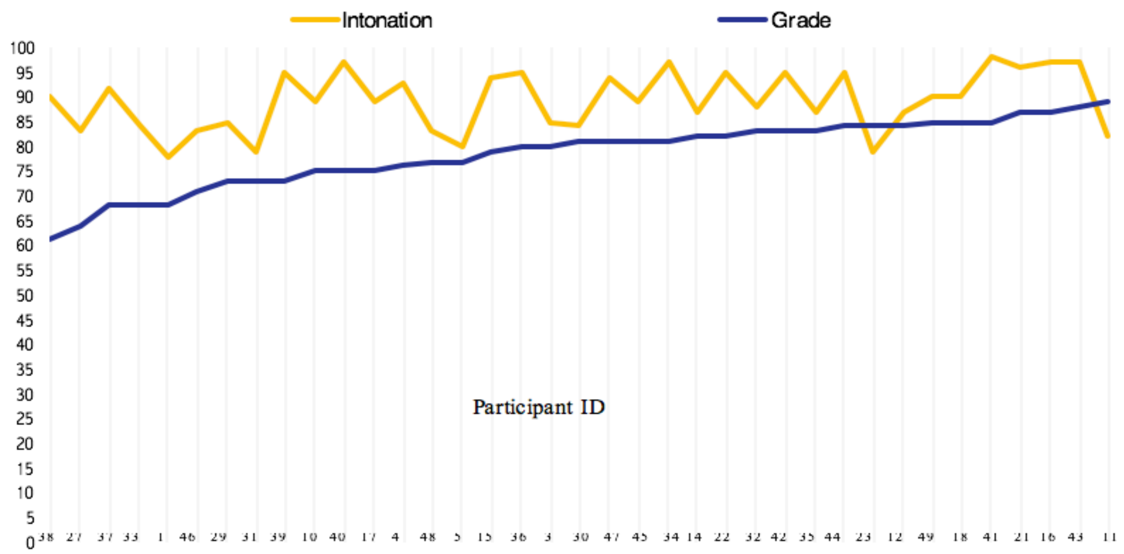


Figure 47. Graphic Representation – Intonation

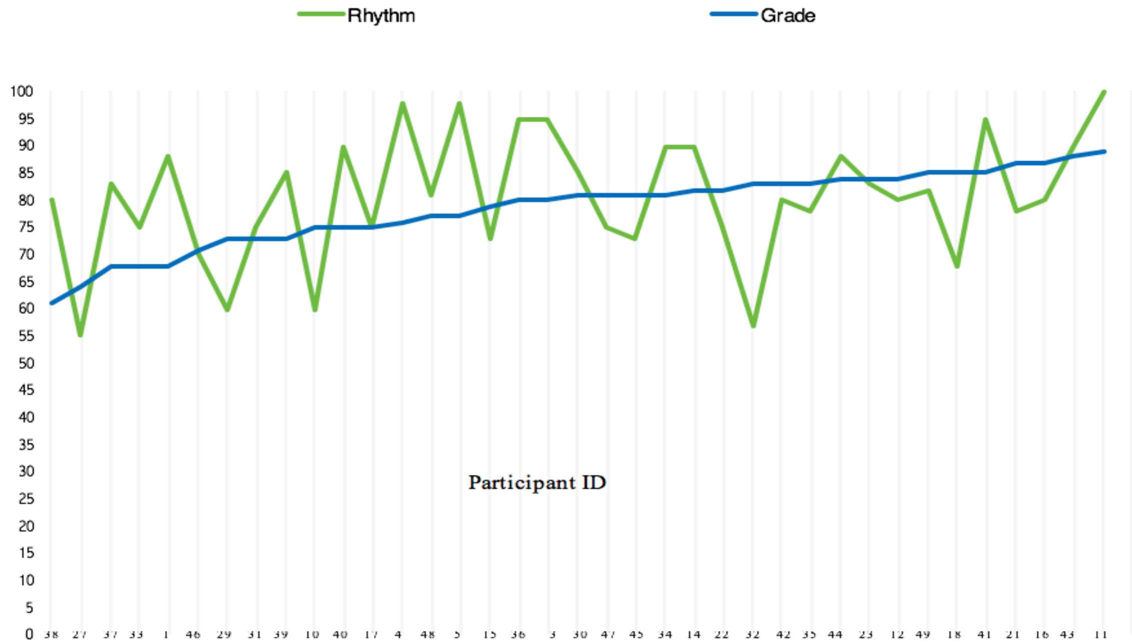


Figure 48. Graphic Representation – Rhythm

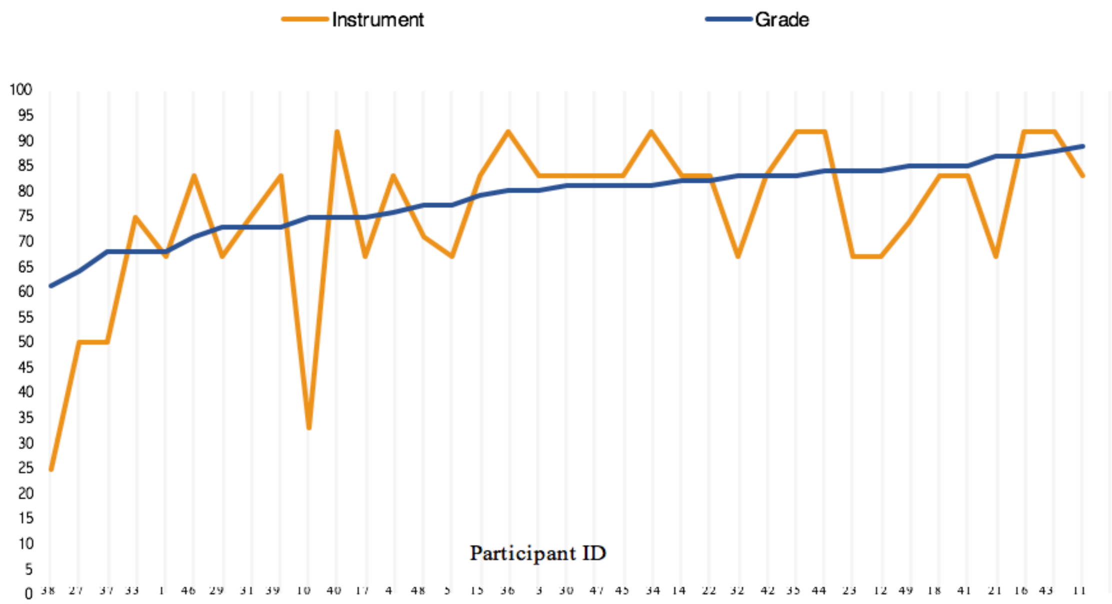


Figure 49. Graphic Representation – Instrument Angle

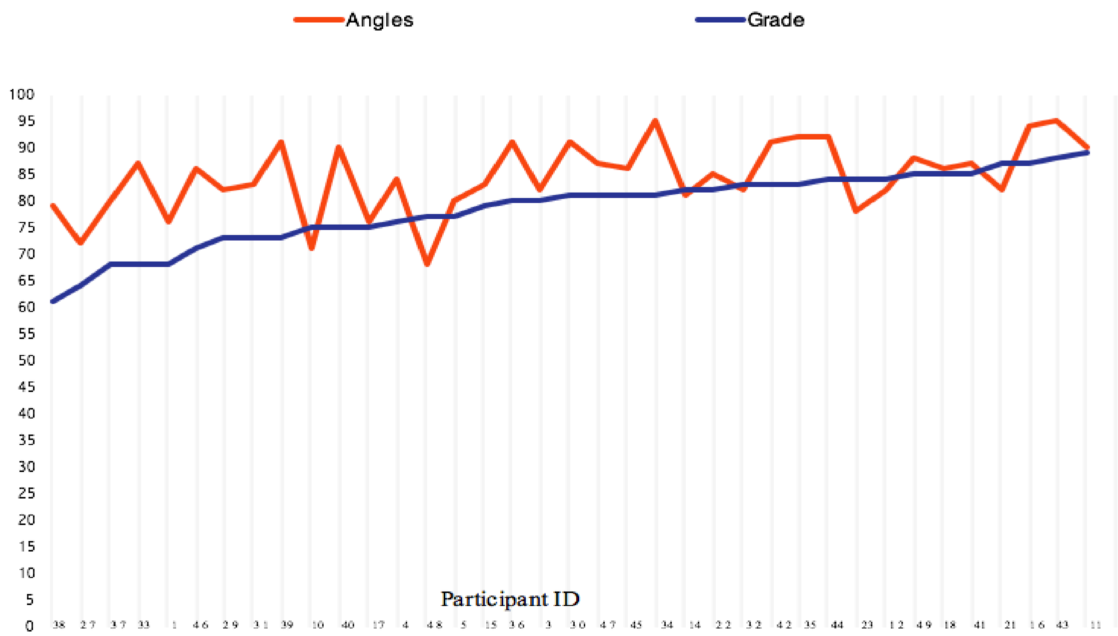


Figure 50. Graphic Representation – Bow Trajectory Angles

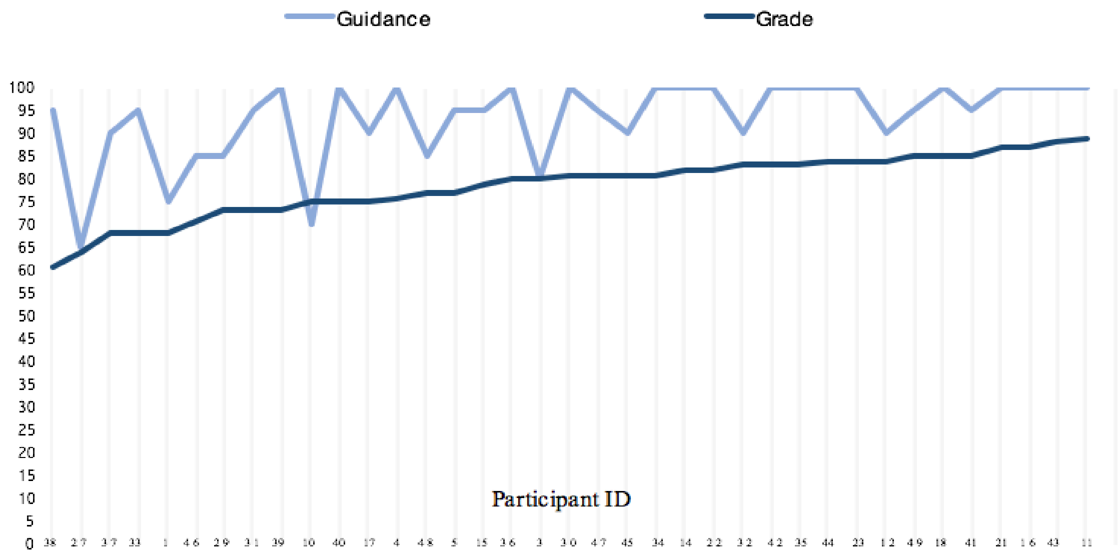


Figure 51. Graphic Representation – Guidance

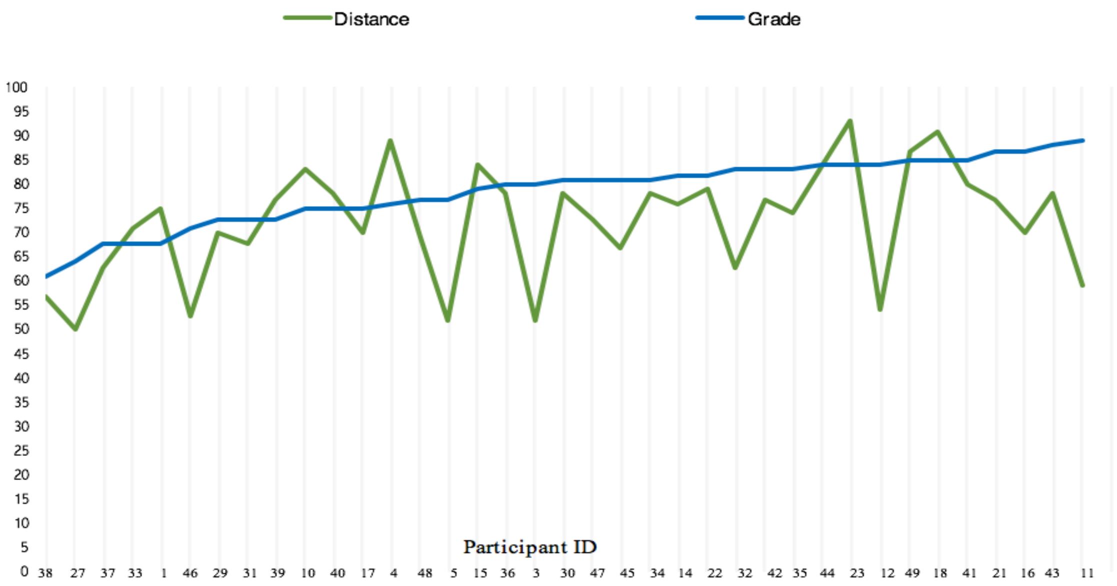


Figure 52. Graphic Representation – Distance

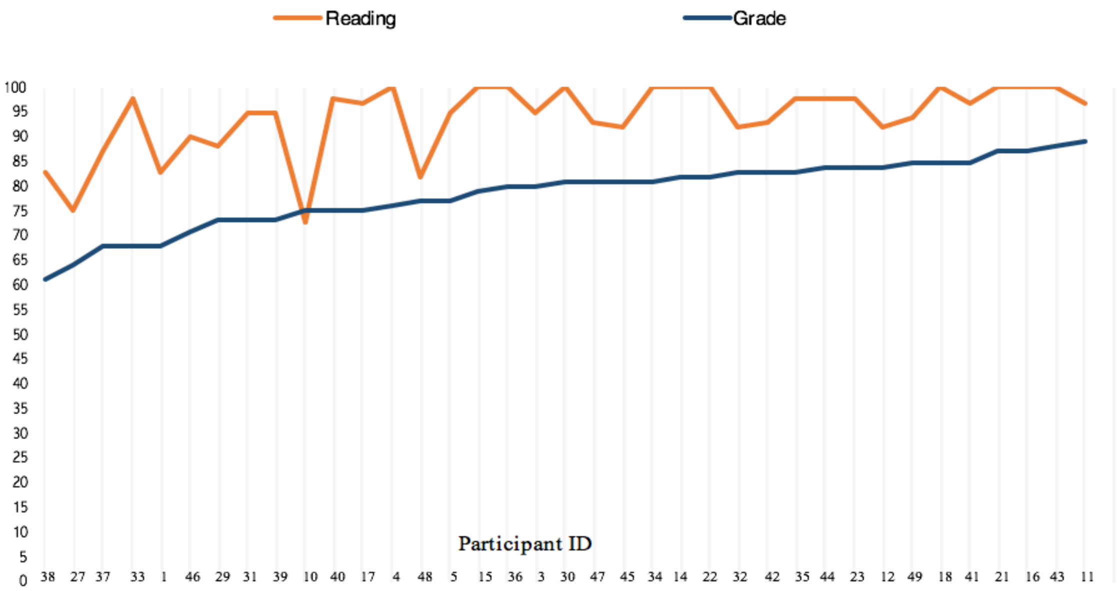
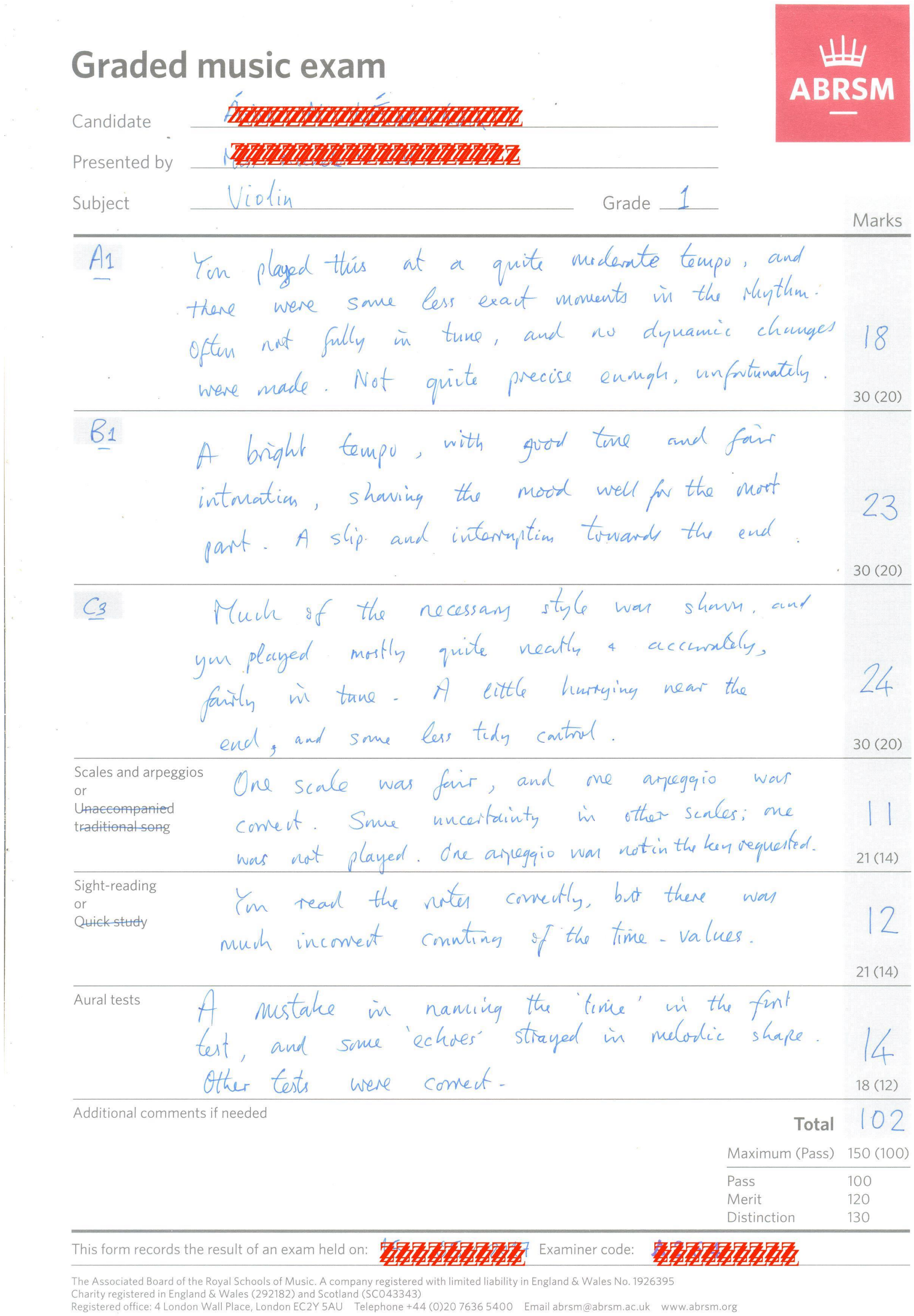


Figure 53. Graphic Representation – Reading

4.10 Students Individual Scores – ABRSM and POP Feedback

The following section illustrates the Primary Observation Package POP feedback sheet and ABRSM examiner remarks sheet for comparison, with respect to 3 students in the study. As may be recalled, the overall mark for an ABSRM grade is awarded out of 150, which is made up as follows: For each of three pieces played, students are awarded up to a maximum 30 marks. For both scales/arpeggios presented, and their rendition of a sight reading test, students are awarded up to a maximum of 21 marks, Finally, for aural responses to four extracts played by the examiner on the piano, students are awarded up to a maximum of 18 marks. Hence, the total score is derived as: score for Piece 1 (30 marks max.); score for Piece 2 (30 marks max.); score for Piece 3 (30 marks max.); score for Scales/arpeggios (21 marks max.); score for Sight-reading (21 marks max.); score for Aural tests (18 marks max.) = Overall Score (150 marks max.). For the purposes of this study, the scores were then normalised out of 100%, such that (for example) a student's ABRSM grade of 75 out of 150 would be converted to 50%.

In Figure 54 below, we have the ABRSM grading report for Student 1. As can be seen, for each of the three pieces played (a, b, c) the student was awarded 18 out of 30, 23 out of 30, and 24 out of 30. For the Arpeggios, the student was awarded 11 out of 21, and was awarded 12 out of 21 for reading. Finally, the student was awarded 14 out of 18 for the aural test. Hence the student was awarded a total of 102 out of 150 for his ABRSM grading. When this is converted into a percentage value for the purposes of the analysis, the value obtained is 68, which is shown in the right hand column of the POP scores in Figure 55. The comments made by the ABRSM rater are then considered in the light of an analysis of the student's performance, as considered in the light of the information gathered by the objective data capture hard- and soft-ware. It is evident that these different modes of appraisal highlight very different aspects of the pupil's playing and overall performances. Repeating the exercise with students 5 and 10 reveals similar discrepancies between the opinions and emphases of the ABRSM rates, when compared with the objective data-driven technical appraisal.



Graded music exam

ABRSM

Candidate [redacted]

Presented by [redacted]

Subject Violin Grade 1

		Marks
A1	You played this at a quite moderate tempo, and there were some less exact moments in the rhythm. Often not fully in tune, and no dynamic changes were made. Not quite precise enough, unfortunately.	18 30 (20)
B1	A bright tempo, with good tone and fair intonation, showing the mood well for the most part. A slip and interruption towards the end.	23 30 (20)
C3	Much of the necessary style was shown, and you played mostly quite neatly & accurately, fairly in tune. A little hurrying near the end, and some less tidy control.	24 30 (20)
Scales and arpeggios or ~~Unaccompanied traditional song~~	One scale was fair, and one arpeggio was correct. Some uncertainty in other scales; one was not played. One arpeggio was not in the key requested.	11 21 (14)
Sight-reading or ~~Quick study~~	You read the notes correctly, but there was much incorrect counting of the time-values.	12 21 (14)
Aural tests	A mistake in naming the 'time' in the first test, and some 'echoes' strayed in melodic shape. Other tests were correct.	14 18 (12)

Additional comments if needed

Total 102

Maximum (Pass)	150 (100)
Pass	100
Merit	120
Distinction	130

This form records the result of an exam held on: [redacted] Examiner code: [redacted]

The Associated Board of the Royal Schools of Music. A company registered with limited liability in England & Wales No. 1926395
Charity registered in England & Wales (292182) and Scotland (SC043343)
Registered office: 4 London Wall Place, London EC2Y 5AU Telephone +44 (0)20 7636 5400 Email abrsm@abrsm.ac.uk www.abrsm.org

Figure 54. ABRSM Example Grading Student 1

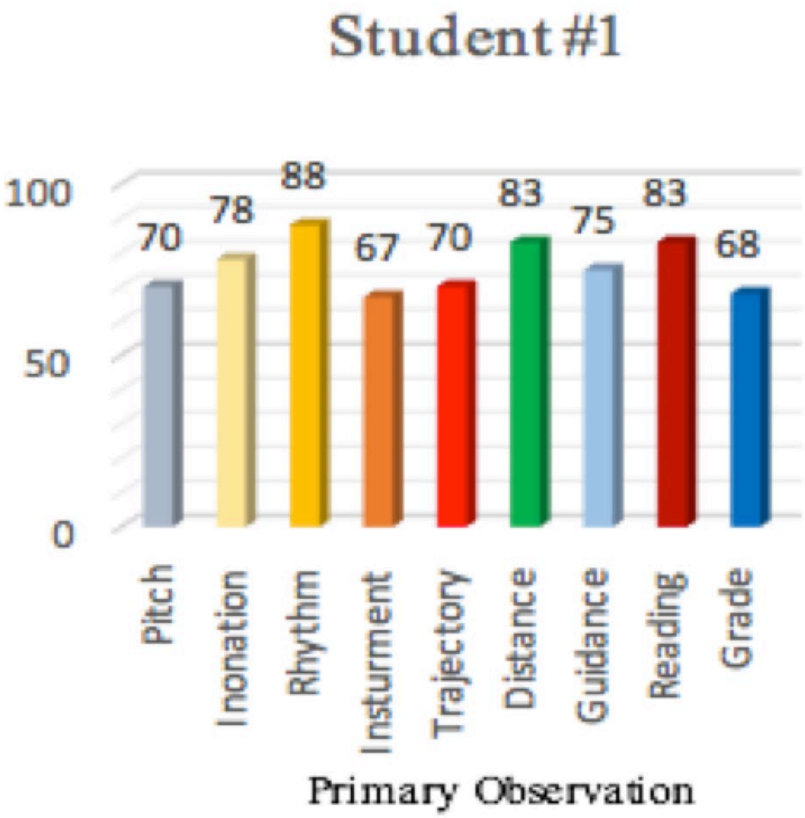
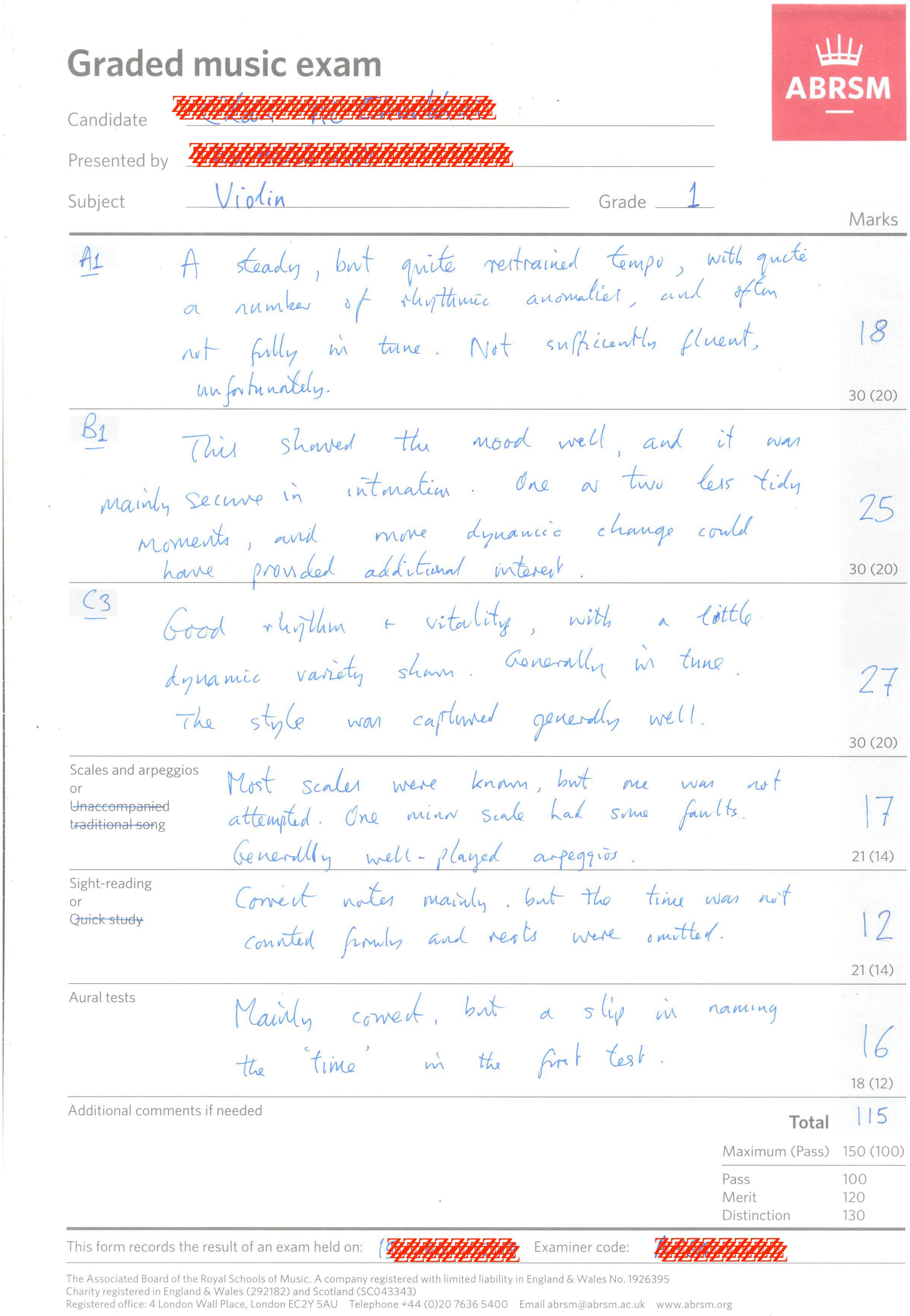


Figure 55. POP Scores Student 1

Student 1 POPs Objective Assessment	Student 1 ABRSM Subjective Appraisal
The teacher, when looking at the Primary Observation Package data for this student should firstly be drawn to the angle at which the instrument is held. The accompanying problems which this posture trait causes such as extra pressure on both hands to hold the bow in place rather than letting it sit on the strings, compromised tone production with associated bow angle trajectory discrepancies and left-hand posture problem associated with the instrument being held too low, which will make it difficult for the student to change positions at a later stage of learning. The next point of interest is the capacity of the student to distinguish between pitches which are in tune with the tonality of a piece and ones which are not. In order for meaningful progress to be made by the student this aspect should be followed up directly as this fundamental learning milestone will impact on the ability of the teacher to instil a sense of intonation. Musicality will be slow to gain traction without this difficulty being addressed. These two points will have the strongest influence on the student's grade outcome and are the main issues behind her current learning gaps with this instrument, to be addressed.	Piece (a) You played this at a quite moderate tempo and there were some less exact moments in the rhythm. Often not fully in tune and no dynamic changes were made. Not quite precise enough, unfortunately. 18/30. Piece (b) A bright tempo, with good time and fair intonation, sharing the mood well for the most part. A slip and interruptions towards the end. 23/30. Piece (c) Much of the necessary style was shown and you played mostly quite neatly and accurately, fairly in tune. A little hurrying near the end, and some less tidy control. 24/30 One scale was fair and one arpeggio was correct. Some uncertainty in other scales, one was not played. One arpeggio was not in the key requested. 11/21. You read the notes correctly, but there was much incorrect accounting of the time values. 12/21. A mistake in naming the time in the first test and some echoes strayed in melodic shape. Other tests were correct. 14/18. Total Score = 102



Graded music exam

ABRSM

Candidate [redacted]

Presented by [redacted]

Subject Violin Grade 1

		Marks
A1	A steady, but quite restrained tempo, with quite a number of rhythmic anomalies, and often not fully in tune. Not sufficiently fluent, unfortunately.	18 30 (20)
B1	This showed the mood well, and it was mainly secure in intonation. One or two less tidy moments, and more dynamic change could have provided additional interest.	25 30 (20)
C3	Good rhythm + vitality, with a little dynamic variety shown. Generally in tune. The style was captured generally well.	27 30 (20)
Scales and arpeggios or ~~Unaccompanied traditional song~~	Most scales were known, but one was not attempted. One minor scale had some faults. Generally well-played arpeggios.	17 21 (14)
Sight-reading or ~~Quick study~~	Correct notes mainly, but the time was not counted firmly and rests were omitted.	12 21 (14)
Aural tests	Mainly correct, but a slip in naming the 'time' in the first test.	16 18 (12)

Additional comments if needed

Total	115
Maximum (Pass)	150 (100)
Pass	100
Merit	120
Distinction	130

This form records the result of an exam held on: [redacted] Examiner code: [redacted]

The Associated Board of the Royal Schools of Music. A company registered with limited liability in England & Wales No. 1926395
Charity registered in England & Wales (292182) and Scotland (SC043343)
Registered office: 4 London Wall Place, London EC2Y 5AU Telephone +44 (0)20 7636 5400 Email abrsm@abrsm.ac.uk www.abrsm.org

Figure 56. ABRSM Example Grading Student 5

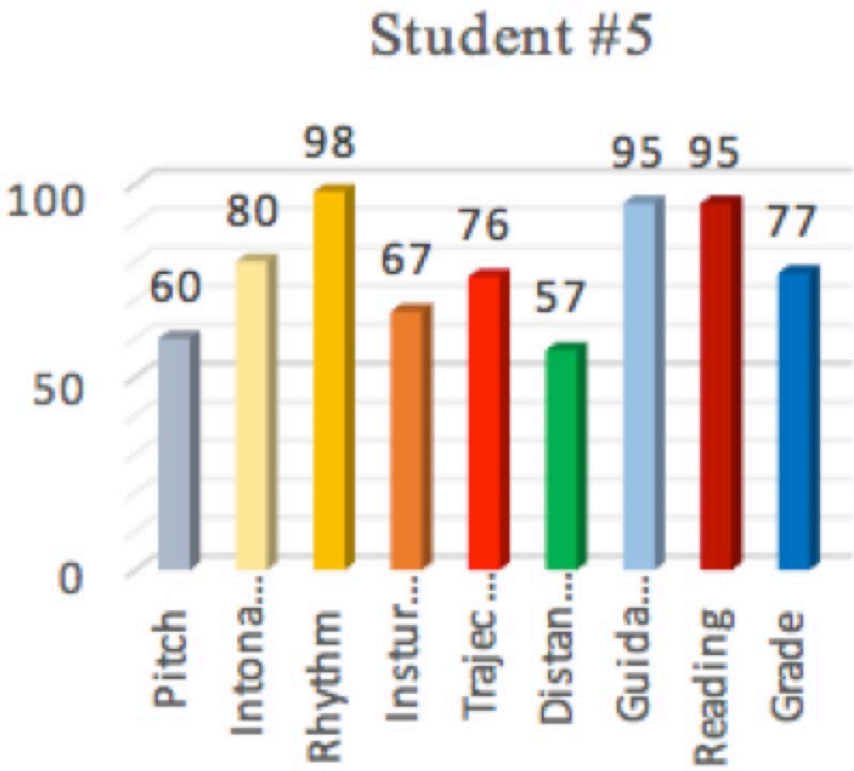


Figure 57. POP Scores Student 5

Student 5 POPs Objective Assessment	Student 5 ABRSM Subjective Appraisal
Again, the teacher is drawn to issues score least well in the observation tool show above. The student does not use enough bow in her playing. This has implications for tone production, dynamic level and sound quality. The posture adopted when holding the instrument and bow trajectory is also weak. In terms of priority, the cultivation of full bows, sensitivity to pitch discrepancies and the angle of the instrument held by the student are key areas where the teacher should focus. The recurring references to the tempo, and dynamics of pieces, suggests that the examiner is working to a predetermined script, making slight adjustments between candidates. As a marking strategy this has limitations, the most obvious one being an inability to specify learning gaps such as the ones above which are preventing the student from improving her playing. Language which uses terms such as mainly, sufficiently, less, little, and some, give feedback which is ambiguous and difficult to direct.	Piece A. A steady, but quite restrained tempo with quite a number of rhythmic anomalies, and often not fully in tune. Not sufficiently fluent, unfortunately. 18/30. Piece B. This showed the mood well, and it was mainly secure in intonation. One or two less tidy moments, and more dynamic change could have provided additional interest. 25/30. Piece C. Good rhythm and vitality with a little dynamic variety shown. Generally, in tune. The style was captured generally well. Most scales were known, but one was not attempted. One minor scale had some faults generally well played arpeggios. 17/21. Correct notes mainly, but the time was not counted firmly and rests were omitted. 12/21. Mainly correct but a slip in naming the “time” in the final test. 15/18 Total Score = 115

Graded music exam

ABRSM

Candidate [redacted]

Presented by [redacted]

Subject Violin Grade 1

		Marks
A2	A quite moderate tempo for this minuet, and there were some less assured moments in the rhythm. You played correct notes, mainly in tune and showed a little of the necessary style.	20 30 (20)
B2	Played at a suitably lively tempo showing the mood generally well. Mainly good intonation but played at the same dynamic level throughout.	23 30 (20)
C3	This was mostly rhythmical and fairly assured, mainly quite well in tune. Some less tidy moments near the end, and some dynamic variety was needed to provide greater interest.	23 30 (20)
Scales and arpeggios or ~~Unaccompanied traditional song~~	Scales had mostly correct notes, but one minor scale had faults, and the flow was often not even. Some uncertainty in one of the arpeggios.	16 21 (14)
Sight-reading or ~~Quick study~~	Accurate notes and you kept going, but often the time-values were not counted exactly.	17 21 (14)
Aural tests	A mistake in naming the 'time' in the first test, and a slip also in part of the final test: (gradual / sudden dynamic change)	14 18 (12)

Additional comments if needed

Total	113
Maximum (Pass)	150 (100)
Pass	100
Merit	120
Distinction	130

This form records the result of an exam held on: [redacted] Examiner code: [redacted]

The Associated Board of the Royal Schools of Music. A company registered with limited liability in England & Wales No. 1926395
Charity registered in England & Wales (292182) and Scotland (SC043343)
Registered office: 4 London Wall Place, London EC2Y 5AU Telephone +44 (0)20 7636 5400 Email abrsm@abrsm.ac.uk www.abrsm.org

Figure 58. ABRSM Example Grading Student 10

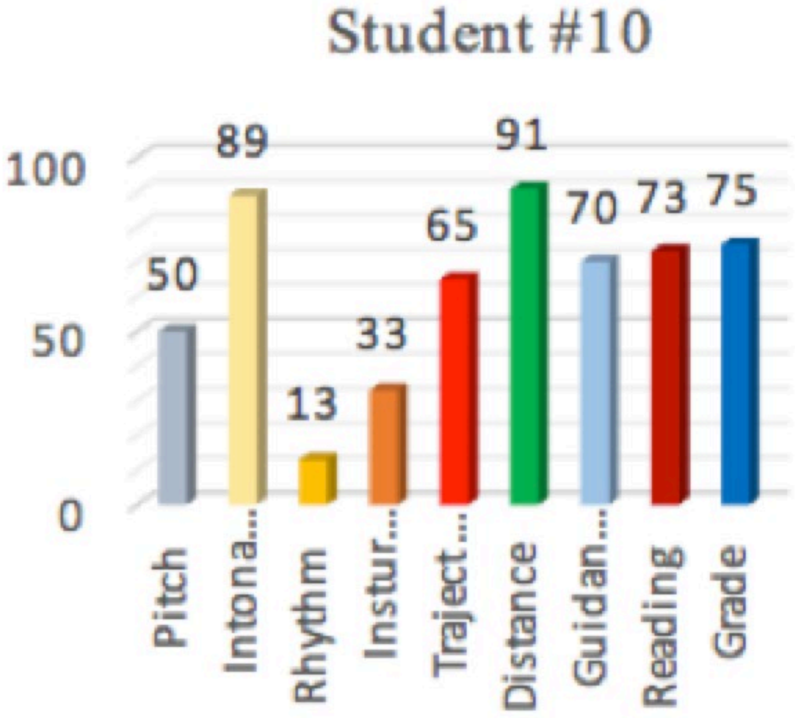


Figure 59. POP Scores Student 10

Student 10 POPs Objective Assessment	Student 10 ABRSM Subjective Appraisal
The observation tool clearly indicates that student 10 has a deficit with rhythm comprehension and implementation. This has been largely overlooked in the subjective grade report. Paragraph 3 furthermore states 'mostly rhythmical and fairly assured', suggesting that the student will be unlikely to register the problem. It is possible to mask an inability in a grade examination through excessive preparation which may compensate the deficit. This point was alluded to by Sloboda (1984, 233). The player uses ample bow but the angle of the instrument should be prioritised as it will compromise the bow trajectory angle. This means that while huge potential is present for tone production, (Distance 91%) poor instrument angle (Instrument 33%) causing poor trajectory, (Trajectory 65%) will have a negative impact on the boy's potential and is likely to produce a compromised yet loud tone. This type of feedback provides a structure for the teacher to remedy the situation. Work could also be done to improve the students understanding of pitch, being culturally and contextuality specific, Time taken with the student to build a sense of tonality would be useful here.	Piece A. A quite moderate tempo for this minuet and there were some less assured moments in the rhythm. You played correct notes, mainly in tune and showed a little of the necessary style. 20/30. Piece B. Played as a suitably lively tempo showing the mood generally well. Mainly good intonation but played at the same dynamic level throughout. 23/30. Piece C. This was mostly rhythmical and fairly assured, mainly quite well in tune. Some less tidy moments near the end and some dynamic variety was needed to provide greater interest. 23/30. Scales had mostly correct notes, one minor scale had faults and the flow was often not even. Some uncertainty in one of the arpeggios. 16/21. Accurate notes and you kept going, but often the time values were not counted correctly. 17/21 A mistake in naming the time in the first test and a slip also in part of the final test (gradual/ sudden dynamic change) 14/18 Total Score = 113

4.11 Discussion

Clearly, the possibility exists that ABRSM examiners are influenced by the way candidates appear to them on the day. Similarly, certain traits which candidates exhibit may have some bearing on the grade outcome. Indeed, the analysis of ABRSM scores by Hargreaves et al. (1998), as reported by Wright (2013, 234), revealed that gender had a significant impact on the grades awarded. It could be argued, however, that musical outcome rubrics which exclude instrument specific criteria run the risk of validity failings when undertaken without these additional insights. Looking at ABRSM in succession it can become apparent what elements an examiner may be fixated on. In the above examples, dynamics and tempo occur continually throughout each student's text suggesting an over reliance on routine or rubric conventions.

To demonstrate the usefulness of the objective approach, it can be seen in Figure 44 on page 157 that participant 10 obtained a score of 13% in the Rhythm variable, 33% in the Instrument variable and 50% in the Pitch discrimination variable. From this we can state clearly that the teacher needs to concentrate on the child's understanding of Rhythm, the way that the instrument is held, and give some consideration to the development of the child's pitch discrimination skills.

One can see from the text of this pupil's ABRSM Grade examination report, two of these points (both non-instrument specific) to be reflective of this but with a slightly ambiguous narrative in relation to rhythm - *some less assured moments in rhythm* yet *mainly in time,* and later, *time values not counted exactly*. In relation to pitch discrimination, this aspect which needs attention is not mentioned and incorrect holding of the instrument is overlooked altogether.

Similarly, the ABRSM grades and comments for the other students 1 and 5 above also indicate that while the rater's comments point out important elements in performance, they are too generic to be of real value to instrumental students with specific learning deficits. Perhaps these observances say more about the differences between teachers and examiners listening preferences, than about individual attributes which students need to address. Also, some examiners and teachers would rather hear a piece of music played in tune at a slow speed, while others would rather hear it played out of tune, but at the correct metronome mark. Clearly, variability in examiner preferences can limit the effectiveness of a test rubric.

In addition, the objective approach can be modified to take account of different observation scenarios such as educational, orchestral, instrumentally specific, learning element specific etc. Being non-intrusive and non-invasive as an approach, it is easily adaptable with minimum disruption to the site of observation or the observation schedule. Technical support needed to analyse the raw data is also minimal, due to the rich profile which the factual measurements uncover.

The posture required for optimum delivery of the mechanical processes is defined and quantified. The sound is analysed for traits which contravene accepted parameters, and the visual analysis makes determinations about the effectiveness of the participant to demonstrate elements of technique, reading ability or other variables deemed to be of significance to the observational objectives from a pedagogical standpoint. This new and current and verifiable profile about the learner, which is the product of the research, is made available to the site of learning in real time, in the spirit of educational Bildung, where transformative learning can take place. This research-based teacher approach to addressing learning gaps is in keeping the views of Hannele Niemi, Professor of Education at the University of Helsinki, who states:

> "Teachers must adopt a research-oriented attitude toward their work. This means learning to take an analytical and open-minded approach to their work, drawing conclusions for the development of education based on different sources of evidence coming from observations and experiences" (Sahlberg, 2015, 117)

The research observations present a new source of information, which is particular to the individual learner, and which provides a working model in which generalist, specialist and student can become directly involved. The evidence based gaps which surface, provide material for a dialogue between pupil and teacher from which a practice regime can be customised. Of course, the science can only take us so far, a qualitative interpretation of the best way forward for any student will obviously be enhanced by the quality of the teacher's methods, or the practitioner's overall strategy and depth of knowledge.

The students who have received this feedback have benefited greatly from the clarity and simplicity (cause and effect) which it demonstrates. More bow – better sound, straight bow – less noise. The minimal technical language required to convey the elements back to children, (coloured blocks for each category) has proved helpful. Very young learners require feedback which is tangible, unsophisticated, and at times simplistic, rather than technologically complex. Sloboda (1985) recommends that 'testing should be carried out with respect to a particular educational question to be answered at a particular time, not to provide a once-for-all statement about capacity or potential for achievement' (p. 234), stating further, that testing should only be used to place people in 'broad categories'. The feedback in this study is complementary to conventional learning methods, is accessible by design to amateur and connoisseur alike, addresses singularities in string learning, is non-judgemental in the way it is returned to the learner and has potential for customisation.

Some students were surprised to discover that they were using so little bow and were only convinced when shown the video clip to support the data. Others did not realise that the right elbow angle was not changing between nut and tip, necessitating the upper right shoulder to compensate. Another student was surprised to discover that the bow directions were being scrutinised, as she felt that it did not make any difference to the sound. It was explained that, while she may be correct in stating this, the question being asked of her was to implement the bowing indications as they appeared on the page, (element 3) of the reading test variables. The dialogue which the feedback opens is stimulating and thought provoking, and provides clear focus on individual learning gaps to which students can respond.

4.12 Findings

Clearly, there is a significant relationship between musical literacy and performance outcomes. This finding backs up many teachers' assertions that to read music is a necessary element to be pursued in music learning – as would be a literacy requirement in other disciplines – despite trends away from music literacy, in terms of the way modern music is produced and sold in record shops and online without it. Having this relationship between literacy and performance confirmed in the comparison study is both unexpected and reassuring. Similar findings in the general population are likely to be found.

The other main correlations, including that of Instrument ($r = .597$) with grade, indicate how important the visual – as well as practical – implications of holding the instrument correctly really are. The appearance of the instrument being held in a way which demonstrates aesthetic, and mechanical contradictions, would appear to bring about a shift from the subjective to the objective in the examiner. Put simply, expectations are diminished when limitations are demonstrated.

Distance ($r = .589$) was another variable with a strong correlation with grade result, which was less surprising. In the pilot study, it became clear that there was a relationship between bow distance and intonation. The intuitive analysis model found a correlation of $r = .410$ between distance and intonation, and a correlation of $r = .378$ between distance and grade. The elbow angles were the main drivers of relationship strength between trajectory and grade result, accounting for most of the correlation with point elbow angle $r = .490$, and nut elbow angle $r = -.484$, indicating both a strong positive and negative linear relationship for elbow angle at each end of the bow. Guidance ($r = .524$) and Pitch ($r = .382$) had moderate correlations with grade result, as expected.

The correlation between Rhythm and Grade ($r = .097$) was low and it is postulated that the construct adopted to measure this variable was weak in design. The bow trajectory variable 5 which contained 6 components (3 relating to bow and 3 relating to elbow angles) isolated 'point' and 'nut' elbow angles as the active predictor of correlation. Future research into the relationship between Rhythm and Grade should contain elements which would more clearly distinguish between comprehension and implementation, as the low correlation is counter-intuitive. In this regard, the analysis was inconclusive, owing to the multifaceted nature of the rhythm attribute data.

The null hypothesis assumes that "an obtained sample distribution can be described by a particular parent distribution" (Hughes and Hase, 2010, 101). A statistical test should consist of a null hypothesis, and tools to test the compatibility of sample data. "The smaller the *P*-value computed from sample data, the stronger the evidence against H_0." (Brase and Brase, 2017). *P*-value (probability of chance) describes the probability that the results of the statistical test are due to chance.

The probability of rejecting the null hypothesis incorrectly is reflected in the *P*- value adopted (α is a level of significance of probability, set for rejecting the H_0 when it is, in fact, true). *P*-values "imply a greater confidence that the apparent difference seen between groups is a reflection of the samples not coming from the same parent population" (Keller, 2006, 101). The hypothesis test has been set at the .05% significance level and where higher levels of significance are present at the .01% level, they are reported thereafter.

The analysis of the data presented here indicates moderate to significant correlations between seven of the eight exogenous observation variables to the single endogenous grade outcome variable. The highest correlation coefficient, r = .656 significant at .01%, was recorded between the reading exogenous variable and grade outcome. A regression prediction equation of $y = 22.402 + 0.5978x$, $R^2 = 0.37629$ was also recorded for the reading variable. Analysis was utilised to explore the latent exogenous variable interplay described by Casella & Berger (2017), and it was found that the point and nut bow angles component mostly affected the exogenous bow trajectory variable.

A high correlation (r = .746 significant at .01% level) was recorded between exogenous variables – Bow Guidance and Bow Distance. The strength of relationship, prediction equations and underlying latent variables, respectively, contribute to a model which demonstrates the importance of music literacy as the highest predictor of grade outcome. Graphic representations of relationships present, make associations explicit and highlight key learning impediments. These representations relate directly to participant ID and generate much needed formative feedback for use in the teaching context, widening the remit of the research.

4.13 Conclusion

Based on these findings, the null hypothesis is rejected, and the alternative hypothesis adopted. The association between the exogenous variables and endogenous variable is moderate to strong for seven of the predictors, and moderate to weak for the remaining exogenous variable. Associations between exogenous variables only are moderate to high. The sample n = 80 was reduced to n = 37, as many of the participants did not wish to fulfil all the participation criteria, namely undertaking grade examinations on which the comparison could be based, despite being capable of doing so.

Having established this association, regression analysis has shown how predictions can be made about response variable scores, with moderate to high levels of accuracy. Causal path analysis has established many of the latent variable components which give rise to the explanatory power of the models put forward. The Intuitive models have shown R^2 = .404 for dependent variables: Pitch, Distance and Trajectory anticipated to have a bearing on independent variable: Intonation. An R^2 =.427 was recorded for dependent variables: Reading, Instrument and Guidance in relation to their influence on independent variable: Grade result.

The Inclusive model through backward elimination saw moderate reduction in the R^2 value from R^2 =.712 to R^2 =.679 as the weaker dependent variables were eliminated. The three Parsimonious models put forward relating to: Bowing, Reading, and remaining variables which were not eliminated in the regression procedure, show final R^2 values of: R^2 =.392, R^2 =.422, and R^2 =.641 respectively. Graphical representations of the data as it emerged from the sample, provide concrete feedback which can be acted upon in a string learning environment.

5 Chapter Five Conclusion

5.1 Introduction

The concluding chapter commences with an overview of the thesis, assessing the answers to the main research questions, and outlining the implications and contributions of the main findings. Following on from this, the chapter evaluates the limitations of the thesis, provides some recommendations on the basis of the findings, and outlines some possible future research opportunities.

5.2 Overview of the Dissertation

The main objective of the thesis is to identify, capture and quantify objective audio-visual data on the key technical physiological parameters (e.g. bowing posture) and specific competences (e.g. sight reading) which contribute to the music ability of primary school student violinists, to enable them to play at varying grades of excellence. Statistical configurations of this data were then aligned to the subjective grades awarded by an ABRSM examiner, undertaking an assessment of the same students' musical performances, on the basis of personal, instrument generic, connoisseurship.

The first step in conducting this study was to review the relevant academic literature on string learning and assessment, musical theory and educational philosophy. The literature review examined the history of musical assessment and addressed the limitations of the current method of grading musical expertise, as is exemplified by the ABSRM's evaluation of performance approach, which provides for little, if any, concrete feedback to enable student progression and the achievement of Bildung, which the literature demonstrates is of specific importance in instrument playing. The review then addressed the use of emerging technologies in music, showing how ever more sophisticated technologies, and related music information retrieval software, are bridging the gap between music education and computer science and bringing the computer into the musical education classroom. The literature review enabled the identification of hard and software that is appropriate and relevant to answering the research questions, thereby assisting and guiding the researcher in refining the plan of the overall research study, the formulation of the research hypothesis and the derivation of the research questions.

The thesis has three overarching research questions. The first research question was: Does an instrument specific approach to performance assessment correlate with musical outcomes in practical ABRSM grade examinations? The analysis of the use of the Associated Board of the Royal Schools of Music's rubrics revealed that the grading system, by design, makes determinations not about the child's level of technical mastery of the instrument, during the process of playing (as the examiners are not required to be specialists in the instrument discipline being examined, and would therefore be unable to ascertain, for example, poor bowing technique on the part of the child), but instead adjudges the overall musicality of the performance, based on what the ABRSM describes in its literature as 'musical outcomes'. So this question asks if attributes, deemed to be essential by a music specialist in the instrument discipline being examined, can be measured, scored, and compared to conventional grade outcome results. This question is at the heart of the research and central to the observation study.

Analysis of the correlations between the subjective ABRSM grade, and the data on students' playing technique, gathered by means of new ICTs, reveals that all of the variables (pitch, intonation, etc.) had positive correlations with the ABSRM grade awarded. However, the correlations were relatively low and ranged from .103 (Rhythm) to .598 (Instrument). Hence only 36% (i.e. .598 x .598) of the variation in the Grade variable was accounted for by variations in the Instrument variable. This result suggests that, while there is a correspondence between the ABRSM grade and technical measures of violin playing technique, the link is not as strong as might be expected, given that the ABRSM considers that its examiners' judgements are beyond reproach.

The second research question was: How can educational Bildung inform string learning? The discussion about Bildung in the literature review revealed the concept to be amorphous and problematic. Nevertheless, the review showed that Bildung is a core idea in music education and is relevant in two ways. First, musical Bildung is part of Bildung in general, which supports the development of a cultivated person and thereby concerns non-musical goals such as fostering intelligence or creativity. Second, there is specific Bildung in music, in terms of gaining musical knowledge and skills. Moreover, central to the idea of Bildung is that of improvement and the realisation of personal potential for excellence, by means of self-reflection and improvement.

When it comes to stringed instrument playing, excellence is achieved by means of adjustment following self-reflection on individual action and practice. This process of the realisation of musical virtuosity by self-reflection is only possible where players can see what they are doing wrong. However, the ABRSM grading process does not allow for instrument specific feedback on the technicalities of playing. Hence, although the students' performances receive ABRSM grades, students are given little or no idea as to how they might improve their technique. By contrast, the process of audio-visual data capture, which is at the heart of this study, enables students to examine for themselves, in great detail (and repeatedly, if necessary), which aspects of their instrument playing technique may require remedial action, and apply appropriate corrective measures. Hence in this manner, the process of data capture facilitates the achievement of educational Bildung by the student. Additionally, when students can see the problems with their technique by means of technology, this de-personalises the process of critical appraisal, altering the dynamics between student and teacher, and democratises them, making the teacher more of a mentor than a critic, and helping to realise Bildung for both.

The final research question is: To what extent can an objective observation process predict a subjective 'musical outcome' product? To answer this question, recourse was made to bi-variate and multi-variate regression analysis. Firstly, bi-variate regression was used, and eight equations (from Y = α(constant) + b1 X1 *to* Y = α(constant) + b8 X8) were calculated. In this instance Y was the ABRSM grade achieved, while X_1 to X_8 represents the independent variables, from pitch to reading. The R^2 values of the resultant equations ranged from 0.011 to 0.37; thus 37% of the variance in ABSRM grade was explained. Hence, as with the correlation analysis, the bi-variate analysis did not provide overwhelming validation of the subjective ABRSM grade when considered in the light of the objective measures of various technical aspects of violin playing.

The final regression analysis used a multi-variate approach in which Y (Grade achieved) = α(constant) + b1 X1 (Pitch) + b2 X2 (Intonation) + b3 X3 (Rhythm) + b4

X4 (Instrument) + b5 X5 (Trajectory) + b6 X6 (Distance) + b7 X7 (Guidance) + b8 X8 (Reading). The regression models chosen for this study included the intuitive (based on the researcher's skills and knowledge of the process of teaching the violin), the inclusive (which maximises the number of independent variables) and the parsimonious (which maximises the R^2 value of the model). The R^2 values for these models were Intuitive = .427; Inclusive = .712; Parsimonious = .679. These results show that when all the objective variables are used in unison to predict the ABRSM grade, they can account for a creditable 70% of the variance in the grade. These results can also enable a testing of the study's domain hypothesis, which was: Relationships between objective observation scores and subjective grade results are random. H0 $p=0.5$. The alternative hypothesis asserts that there is a link between the two. H1 $p\neq0.5$. The regression analysis indicates that the hypothesis is rejected, and the alternative hypothesis is supported. Moreover, comparing the results for bi-variate and multi-variate analysis suggests that single elements of violin playing expertise (pitch, intonation, etc.) that have been measured, are less individually important than the fact that they mesh together and bolster each other. Thus, where one of the mutually supportive elements (say, posture) falters, it necessarily undermines the others (such as bowing), and thereby weakens the overall musical technique.

5.3 Key Contributions and Implications

The thesis provides a significant contribution to the body of knowledge on the process of assessing students' mastery of violin playing (although few such studies have been undertaken), and demonstrates the limitations of the current wide spread use of subjective appraisal based on generic connoisseurship, and provides the foundations for a fairer, unbiased, objective grade calculation. Additionally, it demonstrates how new, relatively inexpensive, mobile technologies can be used in musical education to enable students to quickly and easily obtain explicit, detailed (but readily comprehensible), feedback on a variety of parameters that are considered critical to the achievement of playing excellence, and thereby take remedial action in order to improve musical performance. Furthermore, such a new approach to feedback in musical education both champions and encourages the development of music Bildung, for both students and their teachers.

In addition, in chapter 2, the utilities of task specific holistic and analytical rubrics for the assessment of musical performance were examined, and their strengths and weaknesses considered. Wesolowski's (2012) study has shown that even analytic rubrics (although better able than their holistic counterparts) have limited utility in providing assessment information about a student's ability to complete primary tasks within music literacy, such as reading written note pitches. These deficiencies in rubrics led Wesolowski to examine in more depth the critical role of expert raters (connoisseurs) in musical performance assessment. Following a large scale (N=1704) analysis of 142 musical raters, Wesolowski endorsed the approach adopted for the thesis by stating:

> As the field of music education, and arts in general, becomes more reliant upon data-driven evidence of student achievement and program effectiveness, it will surely be looking more toward the fields of educational measurement and data science to provide insightful methodologies to both improve the validity, reliability, and fairness of music assessment contexts and as a means to discover new empirical patterns underscoring music teaching and learning.
> (Wesolowski, 2019, 622)

As well as these key contributions, the research offers secondary benefits, that are worthy of note. First, incorrect posture while violin playing can lead to serious long term health problems detailed by Aránguiz et al. (2015) and Regenspurger & Seidel (2015). In the case of string playing, the child's hands grow around the instrument as the child grows, and posture cultivated as a child dictates the scope and possibility for the emerging player. The relatively new field of music medicine has identified a growing number of serious musculo-skeletal problems, nerve entrapments and focal dystonia (abnormal tissue tone) according to Aránguiz et al. (2015) with Blanco-Piñeiro et al. (2015) attributing this to incorrect posture and technique being adopted when learning. Two new journals, *Problems of Performing Artists* and *Medical Problems of Performing Artists*, reflect the growing awareness of this problem. By using new technologies to identify incorrect postures, they can be remedied and the associated health problems avoided.

Second, the results obtained by assessing different aspects of instrument playing can enable orchestral directors to make informed decisions about who is going to occupy key desks in each section within the orchestra. For example, a musician who demonstrates a high level of rhythm accuracy, but a low score for other variables, would be more suited to percussive tasks rather than melodic ones. Similarly, students with good posture (as evidenced by pitch angle) should be selected to play at the front desks usually reserved for more accomplished players where leadership can be given, rather than the back desks. Furthermore, musicians scoring highly on the reading variables would be suitable candidates to be selected for lead or solo tasks. Such musicians playing solo material will also need to have strong tone (as evidenced by bowing distance), along with high scores in many of the other variables, (such as intonation, bow angles and reading).

The assignment of musicians to roles within an orchestra is frequently problematic. However, the use of objective data, obtained through the use of new technologies, can help to explain to students and professionals why certain players are chosen over others, for certain tasks. Where the basis of such decisions is grounded in fact (rather than favour), it can be more readily explained to a student as a matter of course, where the learning gaps exist, permitting a more democratic progression for instrument players, as they progress through the ranks of an orchestra, in much the same way as athletes' progress in sport with the aid of motion capture and sports technology. Additionally, when such decisions about roles given to musicians within the orchestra are made in this way, they should lead to an improved musical performance. Furthermore, the model can be cultivated at a higher level where ensemble directors are tasked with ensuring certain skill sets are present in musicians, governed by the demands of the repertoire. The ways in which the basic model can be modified for different types of observational objectives is considered below.

5.4 Limitations of the Thesis

In retrospect, and with the hindsight of self-reflection, it is apparent that the thesis has some limitations. Firstly, the data gathered and analysed in the study was generated from a single primary musical class in Ireland. Consequently, the findings of the thesis may have limited generalisability to other musical classes, both in Ireland and beyond. However, the techniques used for data collection could readily be used to replicate the study with another class of, for example, secondary school violinists, or musicians using other instruments. The sample size used in the thesis is sufficient such as to suggest that the results are statistically significant, but a larger sample could have made the study more authoritative. However, whether or not children and their parents prefer to avoid taking the ABRSM examination is a personal decision, which is outside the researcher's control. Secondly, the results reported in the thesis are tied to the date and time when the cross-sectional study took place. Hence, the results of the students' violin playing performances are culturally and geographically specific to the context of Ireland at a specific time, which may constitute a limitation to the study's generalizability, although the process of instrument learning for standard repertoire is universally identical, wherever it takes place.

Thirdly, although the technologies used in the study were fit for purpose in terms of the thesis, they were relatively unsophisticated by the standards of this kind of audio-visual data capture. Clearly, the use of more sophisticated equipment, such as the 3d augmented mirror created by Ng et al. (2007) and the MusicJacket developed by van der Linden et al. (2011a), could have provided more accurate information. However, such systems can require considerable lab support, which was not available to the researcher. The continuing development of data capture technologies means that if the same study was undertaken today, it would probably involve even less intrusive equipment, and would have produced more finessed data. Whether such refinements would have produced markedly different results, is a matter of conjecture.

Fourthly, with respect to the statistical methods utilised, similarly there are some limitations. For example, a 5% significance level was used as a "cut-off" point in the statistical tests. However, the choice of a significance level is arbitrary – a 10% level could have been just as readily used, and would have presented (for example) causal paths somewhat different from those displayed in the text. In addition, it would have been possible to use other statistical tests to explore the data in more depth and analyse more fully, possible reasons for differences in pupils' performances. For example, the Chi-Squared test could have been used to test whether gender has any impact on musical excellence, while using ANOVA (One-way analyse of variance) could show whether there were any differences between the musical grades of fee paying private students, and non-fee paying pupils. However, despite such limitations, the study produced some robust results, leading to several recommendations, as detailed in the next section.

5.5 Recommendations

Based on the results of the thesis, several recommendations are offered to music teachers, their students, and policymakers seeking to improve musical training. Firstly, based on the statistical analyses that demonstrated the significant impact of the chosen instrument playing parameters on students' grades, music teachers should be encouraged to avail themselves of similar data capture hard and software, in order to enrich the feedback possibilities available to their pupils who are learning to play the violin. Secondly, following from this, national policymakers responsible for enhancing the quality of music education in Ireland, should consider running a pilot scheme in Mayo and providing the necessary data capture technology and also establishing training programmes for music teachers, so that they can learn how to use it in the manner described in the thesis. The objective would be to assess the progress of students in achieving mastery of their instruments, and compare their performance with students who do not have access to such data capture technologies, and therefore do not benefit from the fulsome feedback made available to students by technology enhanced data capture. The overarching long-term aim would be to increase the extent of musical competence and personal motivation among junior musicians at national level and thereby enhance their opportunity to become career musicians. This would also help to raise the level of awareness within the provision of musical education regarding the potential of mobile technologies as teaching aids.

5.6 Future research

The thesis has shown how disparate elements of string playing ability can be observed objectively and measured accurately with the help of new technologies. Such measures could be directly linked to an overall musical outcome. Hence future research could look at ways in which these separate elements can be brought together under one platform, to assist string learners and players to self-diagnose their performance problems and learn remotely. The possibility exists to create a computer programme into which all the measurements could be loaded, in order for an overall score to be automatically calculated. Future work in this direction could involve the use of self-evaluating apps such as MusicWrench, Aural Trainer, MocapViewer and PostureScreen to create a single portable observation package that is not lab dependent. Some work is already taking place to facilitate these developments. For example, Bevilaqua et al. (2007), have developed a prototype using a wireless sensor system.

Furthermore, the observational process adopted in the thesis was used to study primary school players. However, through further work, the process could easily be honed for different levels of grading excellence. Players in novice, amateur and professional orchestras will have different criteria, but it should be possible to set and adjust the parameters to suit the repertoire and pool of players being assessed. In addition, further research would enable variables to be designed, which could observe levels of skill that have their own postural prerequisites. Hence, new variables could be identified and tested for measurement of accuracy, smoothness of shift, precision of intonation in the new position, overall dexterity, progression to other positions and smoothness of return – all qualities familiar to string teachers, orchestral musicians and virtuoso players alike. This type of fine-grained observation was intentionally outside the scope of this research, as it focused primarily on novice level abilities and issues around best learning practice.

In a like fashion, future research could be focused on the measurement of more advanced string techniques, such as the quality of portamento, vibrato, or any number of stylistic proclivities. There are six basic types of bowing: détaché, legato, martelé, staccato, spiccato and ricochet which all contain skill sets which could be scrutinised for quality and measured with a bowing sub-variable to study players undertaking intermediate and advanced performances. In short, the technology can be used to measure different playing characteristics and attributes, in accordance with whatever it is that the teacher is trying to teach.

The thesis focuses on the measurement of string playing, but further research could see whether the techniques developed for this study could be used in the education of players of different instruments. Clearly re-engineering the approach to study players of other stringed instruments (viola, cello, etc.) should be relatively easy. However, additional work could focus on trying to transfer the approach to the playing of other, more disparate, instruments, such as keyboard (piano, harpsichord, organ), wind (flute, oboe, bassoon), brass (trumpet, trombone, french horn) and percussion (xylophone, timpani) instruments. Each instrumental discipline will have a range of particular audio-visual indicators which contribution to musical mastery and performance excellence. Through discussions with music teachers and virtuoso players, it should be possible to identify and agree on core sets of best practice attributes, essential for a good performance. Once identified, new data capture technologies could be used to gather and measure the relevant data. As the sensing technology becomes smaller, more modular and more mainstream, and as its compatibility with smartphone technology becomes more seamless, a transfer of the technique pioneered in this thesis to other instruments should be possible.

Finally, this research has revealed concerns about the objectivity and utility of ABRSM grading. However, the Board has been undertaking examinations for over 130 years, with a subjective methodology which has remained largely unchanged. If possible, future research efforts should attempt to involve the ABRSM in the development of better methods of grading, in which subjective appraisal can be endorsed and refined by means of agreed objective measures. The endorsement of such an approach by the Board would lead to a quantum leap in the process of musical performance assessment.

5.7 Dissemination

The dissemination of the results of the thesis is an important element in maximising impact of the research findings. Writing up the thesis is a necessary requirement of a doctoral level award, but is just the first step in the broader process of disseminating the research results. Hood (2002, 3) conceives of the process of dissemination as the "gap-filler" between academic research and its broader application, both in academia and beyond. Hence it is essential to bridge this gap, by dissemination, in order to transfer knowledge from academic researchers to practitioners and policy makers. Academic research is pursued in order to produce additional and original contributions to the existing body of knowledge. Hence, disseminating these contributions, once the thesis is written, is the responsibility of the researcher. However, since the growth in the use of new information and communication technologies, and the virtual infrastructures (like the Internet and the Wide Wide Web) upon which they rely, the process of the disseminating new knowledge has changed radically. Doctoral theses are now made available, via open access policies, for all to read on-line, and become more efficient and convenient. Such additions however small in size, all help to build the global Knowledge Economy and the Information Society. At a local institutional level, printed hard copies and also digital versions of the thesis will be available in the library and the digital research repository http://eprints.lincoln.ac.uk) of the University. At a national level, a digital copy of the dissertation will be deposited in the Digital Repository of Ireland (www.dri.ie). Moreover, the researcher intends to disseminate the findings in a national conference in Ireland, possibly the annual conference of the Educational Studies Association of Ireland. Similarly, at international level, to disseminate the findings of the thesis, it is anticipated that a research paper will be written, based on the thesis, and submitted for publication in a peer-reviewed journal.

In sum, the dissemination strategy will utilise traditional and technical avenues to make the findings of the thesis known to audiences of both academics and musical teaching practitioners. Hence, various steps will be taken to raise the awareness of musical teachers and students regarding the potential of using new data capture technologies in musical education using different channels of communication. These channels will include publications (both printed and on-line), academic and practitioners' conferences, websites and social media to ensure that the findings disseminated as widely and broadly as possible, in the nope of maximising the impact of the research.

5.8 Conclusion

As was described in chapter 1, the main motivation for undertaking this thesis arose from a desire to improve the process of assessment of young violinists. Working as a musical teacher for over 30 years, and preparing countless students for the ABRSM grading examination, it became evident that, where performances did not reach standards that were required by the Board's examiners, the grading process did not provide students with adequate feedback to enable them to improve. This deficiency in the grading process prompted an investigation of the relationship between objective and subjective assessment of novice string players. The use of new data capture technologies has enabled a comparison between summative subjective grade appraisals and formative objective analyses of critical aspects of the process of playing the violin. In consequence, a formative feedback component has emerged with the potential to transform the way performance music education is delivered. It seems evident that, in line with other aspects of human activities, emerging advances in the development of low cost micro sensing technology, coupled with the growth in open source platforms, will play an important part in transforming the way music education is delivered in the future.

The thesis demonstrates that the use of data capture technologies enables immediate appraisals of all the relevant physiological aspects of playing a stringed instrument. Mapping and measuring variables which plot pitch discrimination, intonation accuracy, rhythmic precision, posture angles, tone production and sight-reading abilities, enables pupil players to easily and readily see where they are making mistakes, and also to take remedial action and map their subsequent improvement. As well as aiding the process of improvement, the continuous process of measurement, reflection and change, re-measurement, reflection and more change, enables the student to better understand and perceive the process for self-improvement, thereby instilling a high sense of morale, and helping to build Bildung. Additionally, this process empowers pre-service generalist trainee teachers, who may only possess limited knowledge of string playing, to return accurate feedback directly to the student. The ability to provide accurate, detailed feedback also assists the specialist to open new conversations with the student about gaps in the learning process with which the teacher is already familiar.

In the first chapter, it was mentioned that the Primary School Curriculum, (NCCA, 1999) recommends that *all* subjects be assessed. Moreover, the National Council for Curriculum and Assessment states "Observation helps the teacher to find out the varying degrees of success with which a child acquires and masters different skills and knowledge and then to adjust teaching and learning contexts accordingly" (NCCA, 2007, 46). The thesis demonstrates a process which enables this aspiration to finally be realised in music education, albeit 20 years after these recommendations were first published. The recommendations of the thesis, and the identification of future possible work, demonstrate that although the thesis is complete, the task of incorporating new technologies into musical education has only just begun. Hopefully this initial contribution to knowledge in this area, as provided by the thesis, and its wider application in music education, will be of value to both aspiring young violinists and their teachers alike.

Bibliography

Abdallah, S. A., Plumbley, M. D., (2006). Unsupervised analysis of polyphonic music using sparse coding. *IEEE Transactions on Neural Networks*, *17*(1), 179-196.

Abeles, H. F., (1973). Development and validation of a clarinet performance adjudication scale. *Journal of Research in Music Education*, *21*(3), 246-255.

Abeßer, J., Hasselhorn, J., Dittmar, C., Lehmann, A., & Grollmisch, S., (2013, October). Automatic quality assessment of vocal and instrumental performances of ninth-grade and tenth-grade pupils.[Paper presentation] in *Proceedings* of the 10th International Symposium on Computer Music Multidisciplinary Research, Marseille, France (pp. 975–988).

Abrahams, F., (2015). Another Perspective. *Music Educators Journal*, *102*(1), 97-100.

ABRSM Examiner Information. (Undated). Available from https://gb.abrsm.org/fileadmin/user_upload/PDFs/Examiner_role_description.pdf [accessed 21st March 2018].

ABRSM Marking Criteria. (Undated). Available from https://us.abrsm.org/fileadmin/user_upload/PDFs/Marking_criteria_Examiners_2018.pdf [accessed 21st March 2018].

Aguilar, C. E. & Richerme, L. K., (2016). A descriptive study of music teacher educators' beliefs about policy. *Journal of Music Teacher Education*, *25*(2), 37-49.

Alexander, M. L., & Henry, M. L., (2012). The Development of a String Sight-Reading Pitch Skill Hierarchy. *Journal of Research in Music Education*, *60*(2), 201-216.

Allen, S. E., & Duke, R. A., (2013). The Effects of limited, restricted music practice on overnight memory enhancement. UPDATE: *Applications of Research in Music Education*, *32*(1), 67-73.

Altenmüller, E., Gruhn, W., Gordon, E., (ed.) (1997). *Music, the brain, and music learning*. New York: GIA Music.

Alves, A., (2019). The German Tradition of Self-Cultivation (Bildung) and its Historical Meaning, *Educação & Realidade, 44*(2), 1-18.

Aránguiz, R., Chana-Cuevas, P., Alburquerque, D., & Curinao, X., (2015). Focal dystonia in musicians: Phenomenology and musical triggering factors, *Neurología (English Edition), 30*, 270-275.

Araújo, N.C.K. de, Gatto Cardia, M.C., Másculo, F.S. and Gomes Lucena, N.M., (2009). Analysis of the frequency of postural flaws during violin performance. *Medical problems of performing artists*, *24*(3), 108-112.

Asmus, E. P., (1999). Music assessment concepts. MENC: the national association for music education. *Music Educators Journal, 8*(2), 19-24.

Bačić, B., (2015). Open-Source Video Players for Coaches and Sports Scientists. [Paper presentation] in *Proceedings* of the 33rd International Symposium on Biomechanics in Sports, Poitiers, France (pp. 515-518).

Barbancho, C., de la Bandera, A.M., Tardon, L.J., (2009, April). Transcription and expressiveness detection system for violin music. [Paper presentation] in *Proceedings* of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009). Taipei, Taiwan (pp. 189-192).

Barbieri, P., & Mangsen, S., (1991). *Violin intonation: a historical survey*. UK: Oxford University Press.

Barbour, J. M., (1952). Violin intonation in the 18th century. *American Musicological Society, 5*(3) 224-235.

Bates, V. C., (2011). Preparing rural music teachers: reflecting on "shared visions". *Journal of Music Teacher Education, 20*(2), 89.

Battle, J., (2002). *Culture-free self-esteem inventories.* Austin, Texas: Pro-Ed.

Benedict, C. L., (2012). Critical and transformative literacies: music and general education. *Theory Into Practice, 51*(3), 152-158.

Bentley, A., (1966). *Musical ability in children and its measurement*. New York: October House.

Berendsen, H. J., (2011). *A student's guide to data and error analysis*. Cambridge: Cambridge University Pess.

Bergee, M. J., (1988). Use of an objectively constructed rating scale for the evaluation of brass juries. *Missouri Journal of Research in Music Education, 5*(5), 6-25.

Bernstein, B., (1975). *Primary socialisation, language and education*. London: Rutledge & Keegan Paul.

Berry, W.D., (1985) *Multiple regression in practice*. London: Sage.

Bevilacqua, F., Guédy, F., Fléty, E., Leroy, N., Schnell, N., (2007, June). Wireless sensor interface and gesture-follower for music pedagogy. [Paper presentation] in *Proceedings* of the Seventh International Conference on New Interfaces for Musical Expression, New York University Stern School of Business, United States (pp. 124-129).

Blanco-Piñeiro, P., Díaz-Pereira, M., P., & Martínez, A., (2015). Prevention and rehabilitation: observational study: common postural defects among music students. *Journal of Bodywork & Movement Therapies, 19*(3), 565-572.

Bluman, A. G., (2017). *Elementary statistics a step by step approach*. London: McGraw-Hill.

Born, G., and Hesmondhalgh, D., (2000) *Western music and its others: difference, representation, and appropriation in music*. Berkeley: University of California Press.

Branco, M., Branco, G. A., & Brito, A., (2016). Introduction to data collection for performance analysis and their integration in martial arts training. A biomechanical and motor control perspective. *Revista de Artes Marciales Asiáticas (RAMA), 11*(2), 19-21.

Brase, C.H., Brase, C.P., (2017). *Understanding basic statistics*, Boston, Mass.: Cengage Learning.

Brightwell, G.W.E., (2013) Review: The Associated Board of the Royal Schools of Music: A Social and Cultural History by David C.H. Wright, *Notes, 70*(2), 283-286.

British Educational Research Association BERA (2011). *Ethical guidelines for educational research*. Available at https://www.bera.ac.uk/researchers-resources/publications/ethical-guidelines-for-educational-research-2011 [accessed: 8 August 2016].

Broudy, H., (2008). *A Realistic Philosophy of Music Education*, London: Routledge.

Bryman, A., (2001). *Quantitative data analysis with SPSS release 10 for Windows, a guide for social scientists*. London: Routledge.

Burton, S. L., (2015). Making music mine: the development of rhythmic literacy. *Music Education Research, 57*(5), 1-24.

Buss, E., Taylor, C. N., & Leibold, L. J., (2014). Factors affecting sensitivity to frequency change in school-age children and adults. *Journal of Speech, Language & Hearing Research, 57*(5), 1972-1982.

Butke, M. A., (2014). Assessing expressive movement: measuring student learning outcomes in the general music classroom. *General Music Today, 27*(3), 23-27.

Byo, J., Brooks, R., (1994). A comparison of junior high musicians' and music educators' performance evaluations of instrumental music. *Contributions to Music Education, 21*, 26–38.

Çalgan, G., (2015). Kemandan viyolaya geçiş süreci., (Turkish). Social Sciences Review of the Faculty of Sciences & Letters University of Uludag. *Fen Edebiyat Fakültesi Sosyal Bilimler Dergisi, 17*, 1-9.

Cangro, R., (2016). Student Collaboration and Standards-Based Music Learning. UPDATE: *Applications of Research in Music Education, 34*(3), 63-68.

Casella, G., & Berger, R. L., (2017). *Statistical inference*. Delhi: Cengage Learning.

Charles, J., (2010). *Playing technique and violin timbre: detecting bad playing*. Ph.D. dissertation. Dublin Institute of Technology. doi:10.21427/D7HC8P.

Ciorba, C., and Smith, N., (2009) Measurement of Instrumental and Vocal Undergraduate Performance Juries Using a Multidimensional Assessment Rubric, *Journal of Research in Music Education*, *57*(1), 5-15.

Cohen, M. J., (1997). *Children's memory scale*, San Antonio Texas: The Psychological Corporation.

Conover, W. J., (1980). *Practical nonparametric statistics*. New York: Chichester Wiley.

Cooper, A. A., (Earl of Shaftesbury) (1710) *Soliloquy: Or Advice to an Author*, London: John Morpew (accessed via https://archive.org/details/soliloquyoradvi00ofgoog/page/n1 on 4th November 2019).

Cooper, N., (1994). An Exploratory Study in the Measurement of Children's Pitch Discrimination Ability, *Psychology of Music*, *22*(1), 56–62.

Courvoisier, C., (2006). *The techniques of violin playing*. New York, USA: Dover Publications Inc.

Cowles, M., Davis, C., (1982). On the origins of the .05 level of statistical significance, *American Psychologist, 37*(5), 553-558.

Crittenden, P., (1981) Thoughts about Locke's Thoughts about Education, *Journal of Philosophy of Education*, 15(2), 149-160.

Cunningham, J., Aldrich, J., (2012). *Using SPSS*, Los Angeles: Sage.

Darrow, A.-A., Cassidy, J. W., Flowers, P. J., Register, D., Sims, W., Standley, J. M., Menard, E., & Swedberg, O., (2009). Enhancing literacy in the second grade. *UPDATE: Applications of Research in Music Education, 27*(2), 12-26.

De Cheveigne, A., (2005). Pitch perception models. In: Plack C., and Oxenham A., (eds.) *Pitch, neural coding and perception* (pp. 169-233) New York: Springer.

Di Veroli, C., (1991). *Violin Intonation: more on tempered fifths.* UK: Oxford University Press.

Dietze, S., (2013). *Die musikalische audiobearbeitung mit celemonys melodyne editor 2*. München: GRIN Verlag.

Dittmar, C., Cano, E., Abeßer, J., Grollmisch, S., (2012). Music information retrieval meets music education, multimodal music processing. *Dagstuhl Follow-Ups*, *3*, 95-120.

Dougan, K., (2015). Finding the right notes: an observational study of score and recording seeking behaviors of music students. *Journal of Academic Librarianship, 41*, 61-67.

Doyle, A., (2016). MusicWrench app. Available from https://apps.apple.com/us/app/id965939820 [accessed 20th Dec 2018].

Drake, C., (1993). Reproduction of musical rhythms by children, adult musicians, and adult non musicians. *Perception and Psychophysics, 53*(1), 25-33.

Drake, R. M., (1933). The validity and reliability of tests of musical talent. *Journal of Applied Psychology*, *17*(4), 447-458.

Draper, N., and Smith, H., (1966). *Applied Regression Analysis*, New York: John Wiley and Sons.

Duffin, R., (2007). *How equal temperament ruined music (and why you should care*). NY: Norton and Company.

Dumas, E., (2010). Using combos to improve music reading skills. *Canadian Music Educator / Musicien Educateur au Canada, 51*(4), 28-31.

Dunsby, J., (1995). *Performing music: shared concerns,* New York: Oxford.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E., (1995). Increased cortical representation of the fingers of the left hand in string players. *American Society for the Advancement of Science. 270*(5234), 305.

Elliott, D., (1995). *Music Matters*, Oxford: Oxford University Press.

Ellis, A., (1876). On the Sensitiveness of the Ear to Pitch and Change of Pitch in Music. *Proceedings of the Musical Association*, *3*, 1-32.

Elmore, R., (2006). *School reform from the inside out: policy, practice, and performance*. Cambridge, MA: Harvard Education Press.

Engelhard, G., (1996). Evaluating rater accuracy in performance assessments. *Journal of Educational Measurement*, *33*(1), 56-70.

Fancourt, A., Dick, F., & Stewart, L., (2013). Pitch-change detection and pitch-direction discrimination in children. *Psychomusicology: Music, Mind, and Brain, 23*(2), 73-81.

Farnsworth, P. R., (1969). *The social psychology of music*. Ames, Iowa: Iowa State University Press.

Fautley, M., & Colwell, R., (2018). Assessment in the secondary music classroom. In: G.E. McPherson and G.F. Welch (eds.) *Music Learning and Teaching in Infancy, Childhood, and Adolescence* (pp. 257-276). New York: Oxford University Press.

Fehr, R. C., (2014). Want to use the new music standards in your classroom? NAfME can help. *Teaching Music, 22*, 10.

Field, A., (2002). *Discovering statistics using IBM SPSS statistics: and sex and drugs and rock'n' roll*. Fourth Edition. London: Sage.

Fine, P. A., Wise, K. J., Goldemberg, R., & Bravo, A., (2015). Performing musicians’ understanding of the terms 'mental practice' and 'score analysis'. *Psychomusicology: Music, Mind, and Brain, 25*(1), 69-82.

Fischer, S., (2009). Intonation. *Strad, 120*, 74-75.

Fischer, S., (2013). Intonation. *Strad, 124*, 82-83.

Fisher, R., (1925). *Statistical methods for research workers*. Edinburgh: Oliver & Boyd.

Fiske, H. E., (1983) Judging musical performances: method or madness? *Update, 1*(3), 7-10.

Fleisch, C., (1939). *The art of violin playing*, Chicago: Carl Fischer Inc.

Fleischmann, A., (1998). *Sources of Irish traditional music. 1600-1855*. New York: Garland Publishing Inc.

Flores, R. G., Ginsburgh, V.A., (1996). The Queen Elisabeth musical competition: how fair is the final ranking? *The Statistician, 45*(1), 97-104.

Fox, C., (2003). Temperaments, tonalities and microtonalities: an introduction. *Contemporary Music Review, 22*(1-2), 1-2.

Frederickson, M. L., (2010). The national standards for music education. *Music Educators Journal, 97*(2), 44-50.

Fujii, S.,& Schlaug, G., (2013) The Harvard beat assessment test (H-Bat) a battery for assessing beat perception and production and their dissociation. *Frontiers in Human Neuroscience, 771*(7), 1-16.

Fukuda, T., Ikemiya, Y., Itoyama, K., & Yoshii, K., (2015, July) A score-informed piano tutoring system with mistake detection and score simplification. [Paper presentation] in *Proceedings* of the 12th Sound and Music Computing Conference, Maynooth, Ireland (no page numbers).

Galamian, I., (1966*). Contemporary violin techniques; scale and arqeggio exercises with bowing and rhythm patterns*. New York: Galaxy Music.

Gall, M., Gall, J., Borg, W., (2007). *Educational research: An introduction*, Boston, MA: Pearson Education.

Gamble, S., (1988). The elementary classroom teacher: An ally for music education. *Music Educators Journal, 75*(1), 25-28.

Gardner, H., (1973). Children's sensitivity to musical styles. *Behavioural Development, 19*(1), 67-77.

Gaston, E., (1968). *Music in therapy*. Michigan: Macmillan.

Gerber, C., (2011). A comparison of the teaching philosophy and methods of Philip C. Hayden and the Kodály approach. *Kodaly Envoy, 37*(2), 4-11.

Geringer, J. M., Macleod, R. B., & Sasanfar, J. K., (2015). In tune or out of tune: are different instruments and voice heard differently? *Journal of Research in Music Education, 63*(1), 89-101.

Gerle, R., (1983). *The Art of Practicing the Violin*. Great Yarmouth UK: Gilliard.

Gerle, R., (1991). *The art of bowing practice*. Great Britain: Stainer and Bell.

Gingras, B., Honing, H., Peretz, I., Trainor, L. J., & Fisher, S. E., (2015). Defining the biological bases of individual differences in musicality. *Philosophical Transactions of the Royal Society B*, *370*(1664), 20140092.

Goody, J., (1963). The consequences of literacy. *Comparative Studies in Society and History, 5*, 27-68.

Goolsby, T. W., (1995). Portfolio assessment for better evaluation. *Music Educators Journal, 82*(3), 39-44.

Gordon, E., (1965). *Musical aptitude profile manual*. Boston: Haughton Mifflin.

Gordon, E., (1990). *Predictive validity study of AMMA: A one-year longitudinal predictive validity study of the advanced measures of music audiation*. Chicago, IL: GIA Publications.

Gordon, E., (2002). *Rating scales and their uses for evaluating achievement in music performance.* Chicago: GIA.

Gould, E., (2009). Writing trojan horses and war machines: the creative political in music education research. *Educational Philosophy & Theory, 43*(8), 874-887.

Grachten, M., & Krebs, F., (2014). An assessment of learned score features for modeling expressive dynamics in music. *IEEE Transactions on Multimedia, 16*(5), 1211-1218.

Green, L., (2002). *How popular musicians learn*. U.K.: Ashgate Publishing.

Gresham, F. M., & Elliot, S. N., (1990). *Social skills rating system*. Circle Pines, MN.: American Guidance Service.

Grout, B., (2016). Powerful patterns. *Clavier Companion, 8*, 42-47.

Gudmundsdottir, H. R., (2010). Pitch error analysis of young piano students' music reading performances. *International Journal of Music Education, 28*(1), 61-70.

Guzmán-Valdivia, C., Blanco-Ortega, A., Oliver-Salazar, M., & Carrera-Escobedo, J., (2013). Therapeutic motion analysis of lower limbs using Kinovea. *International Journal of Soft Computing and Engineering, 3*(2), 359-365.

Gzibovskis, T., & Marnauza, M., (2012). Development of young adults' fine motor skills when learning to play percussion instruments. *Music Education Research, 14*(3), 365-380.

Hair, F. G., (1975). Descrimination of tonal direction on verbal and nonverbal tasks by first grade children. *Journal of Research in Music Education, 25*(3), 197-210.

Hansen, D., Milligan, S., (2012). Aural skills: at the juncture of research in early reading and music literacy. *Music Educators Journal, 99*(2), 75-80.

Hargreaves, D., (1986). *The Developmental Psychology of Music*. Cambridge: Cambridge University Press.

Hargreaves, D., North, A., Hatcher, R., (1998) *Establishing Benchmarks and Reliability of Assessments in the ABRSM Examinations*, Music Research Group, University of Leicester: unpublished manuscript.

Hash, P. M., (2011). The universal teacher, by J. E. Maddy and T. P. Giddings (1923). *Journal of Research in Music Education, 58*(4), 384-410.

Hayward, C. M., & Gromko, J. E., (2009). Relationships among music sight-reading and technical proficiency, spatial visualization, and aural discrimination. *Journal of Research in Music Education, 57*(1), 26-36.

Heimonen, M., (2014) Bildung and Music Education: A Finnish Perspective, *Philosophy of Music Education Review*, *22*(2), 188-208.

Helmholtz, H. Von, (1863) *Die Lehre von den Tonempfindungen als physiologische Grundlage für die theorie der music* (1st ed.) Braunschweig, Friedrich Vieweg und Sohn.

Helmholtz, H. von and Ellis, A. J., (1895). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. (translation of the 3rd German edition of 1870) London: Longmans, Green and Company.

Heneghan, F., (2002). *MEND music education national debate*. Dublin: DIT. Available from http://www.musicnetwork.ie/content/files/MEND09d.pdf [accessed 22nd May 2016].

Herron, D., (1985). *Deaf ears report on the provision of music education in Irish schools*. Dublin: The Arts Council.

Ho, T. K.-L., Lin, H.-S., Chen, C.-K., & Tsai, J.-L., (2015) Development of a computer-based visualised quantitative learning system for playing violin vibrato. *British Journal of Educational Technology, 46*(1), 71-81.

Hodgson, P., (1935) *Motion Study and Violin Bowing*. UK: Novello & Company, Ltd.

Hoenig, W.G., Lindlahr, S., Schmidt, A., (undated) *Melodine editor1.2.* Available from https://www.scribd.com/document/53705920/Manual-Melodyne-Editor-1-2-English [accessed 10th July 2018].

Holahan, J. M., Saunders, T. C., & Goldberg, M. D., (2000). Tonal cognition in pattern discrimination: evidence from three populations. MENC: *The National Association for Music Education*, *48*(2), 162-177.

Hood, P., (2002) *Perspectives on knowledge utilization in education.* San Francisco, CA: WestEd.

Horlacher, R., (2004) Bildung: A Construction of a History of Education, *Studies in Philosophy and Education*, *23*, 409-426

Horlacher, R., (2012) What is *Bildung?* Or: Why *Pädagogik* cannot get away from the concept of *Bildung* in (eds.) P. Siljander, A, Kivelä and A. Sutinen, *Theories of Bildung and Growth: Connections and Controversies Between Continental Educational Thinking and American Pragmatism* (pp. 135-147). Rotterdam/Boston/Taipei: Sense Publishers.

Horlacher, R., (2016) *The Educated Subject and the German Concept of Bildung: A Comparative Cultural History*, London: Routledge.

Hughes, I.G., Hase, T.P., (2010). *Measurements and their uncertainties.* U.K.: Oxford University Press.

Humboldt, W. von (1995) Theorie der Bildung des Menschen in A. Flitner & K. Giel (eds.) *Wilhelm von Humboldt Werke in Fünf Bänden:1 Schriften zur Anthropologie und Geschichte*, (pp. 234-240). Darmstadt: Wissenschaftliche Buchgesellschaft.

Humboldt, W. von (N.D.) *Gesammelte Schrifdten*, Berlin Königlich Preussischen Akadamie der Wissenschaften (accessed via https://archive.org/details/gesammelteschrif01humbuoft/page/n7, 4th November 2019)

Hutchins, S. M., and Peretz, I., (2012). A frog in your throat or in your ear? searching for the causes of poor singing. *Journal of Experimental Psychology*, *141*(1), 76-97.

Hutka, S., Bidelman, G. M., & Moreno, S., (2015). Pitch expertise is not created equal: cross-domain effects of musicianship and tone language experience on neural and behavioural discrimination of speech and music. *Neuropsychologia, 71*, 52-63.

Imberty, M., (1969). *L'acquisition des structures tonales chez l'enfant*. Paris: Klinksieck.

Iversen, J. R., and Patel, A. D., (2008, August). The beat alignment test (BAT): Surveying beat processing abilities in the general population, in *Proceedings* of the 10th International Conference on Music Perception, and Cognition (ICMPC 10), eds M. Miyazaki, Y. Hiraga, M. Adachi, Y. Nakajima, and M. Tsuzaki, Sapporo, Japan (no page numbers).

Jaccard, J. L., (2014). Intonation begins in kindergarten: the art and science of teaching music acoustically. *Kodaly Envoy, 40*, 60-67.

Jacobi, B. S., (2012). Kodaly, literacy, and the brain: preparing young music students to read pitch on the staff. *General Music Today, 25*(2), 11-18.

Järvinen-Pasley, A., Wallace, G. L., Ramus, F., Happé, F., & Heaton, P., (2008). Enhanced perceptual processing of speech in autism. *Developmental Science, 11*(1), 109-121.

Jensen, J., & Neff, D., (1993). Development of basic auditory discrimination in pre-school children. *Psychological Science, 4*(2), 104-107.

Jentzsch, I., Mkrtchian, A., & Kansal, N., (2014). Improved effectiveness of performance monitoring in amateur instrumental musicians. *Neuropsychologia, 52*, 117–124.

Johnson, P., (1997). Performance as Experience: the problem of assessment criteria, *British Journal of Music Education, 14*(3), 271-2.

Kafol, B. S., Denac, O., Žnidaršič, J., & Zalar, K., (2015). Analysis of music education objectives in learning domains. *Procedia - Social and Behavioral Sciences*, *186*, 95-104.

Kanno, M., (2003). Thoughts on how to play in tune: pitch and intonation. *Contemporary Music Review, 22*(1-2), 35-52.

Kanno, M., (2007). Prescriptive notation: Limits and challenges. *Contemporary Music Review, 26*(2), 231-254.

Karadag, E., (2012). Basic features of structural equation modeling and path analysis with its place and importance in educational research methodology. *Bulgarian Journal of Science and Education Policy, 6*(1), 194-212.

Kaufman, A. S., (2004). *Kaufman brief intelligence test (2nd ed.)*, Circle Pines, MN.: American Guidance Service.

Keller, D.K., (2006). *The tao of statistics a path to understanding (with no math*). London: Sage.

Kertz-Welzel, A., (2016) Sociological Implications of English as an International Language in Music Education, *Action, Criticism and Theory for Music Education, 15*(3), 53-66.

Kertz-Welzel, A., (2017) Revisiting Bildung and its Meaning for International Music Education Policy, in (eds.) P. Schmidt and R. Colwell, *Policy and the Political Life of Music Education* (pp. 107-122). Oxford: OUP.

Kohn, A., (2000). Standardised testing and its victims. *Education Week, 20*, 46 -47.

Kopiez, R., Lee, J. I., (2008). Towards a general model of skills involved in sight reading music. *Music Education Research, 10*, 41–62.

Kwalwasser, J., & Dykema, P., (1930) *Manual of directions, K-D tests*. New York: Fischer.

Lage, G. M., Borém, F., Vieire, M. N., & Barreiros, J. P., (2007). Visual and tactile information in double bass intonation control. *Motor control, 11*(2), 151-165.

Latimer, M., Bergee, M., Cohen, M., (2010) Reliability and Perceived Pedagogical Utility of a Weighted Music Performance Assessment Rubric, *Journal of Research in Music Education, 58*(2), 168-83.

Law, L. N., Zentner, M., (2012). Assessing musical abilities objectively: construction and validation of the profile of music perception skills. *Plos ONE, 7*(12), 1-15.

Lehmann, A. C., (2014). Using admission assessments to predict final grades in a college music program. *Journal of Research in Music Education*, *62*(3), 245-258.

Lehmann, A. C., McArthur, V., (2002). Sight Reading, in Parncutt, R., &. McPherson, G., (eds.) *The science and psychology of music performance: Creative strategies for teaching and learning* (pp. 135-150). London: Oxford University Press.

Locke, J., (1693) *Some Thoughts Concerning Education*, London: A. and J. Churchill, (accessed from https://archive.org/details/somethoughtscon02lockgoog/page/n7 4th November 2019).

Lowery, H., (1926). Cadence and phrase tests of musical talent. *British Journal of Psychology, 17*(2), 111.

Lundin, R., (1949). The development and validation of a set of musical ability tests. *Psychological Monographs, 63*, 1-20.

Luo, Y. J., Su, L., Yang, Y. H., & Chi, T. S. (2015, October). Detection of Common Mistakes in Novice Violin Playing.[Paper presentation] in *Proceedings* of the 16th International Society for Music Information Retrieval Conference, Malaga, Spain, (pp. 316-322).

Lüth, C., (1998) On Wilhelm von Humboldt's Theory of Bildung, *Journal of Curriculum Studies*, *30*(1), 43-59.

Maestre, E., & Ramírez, R., (2010). An approach to predicting bowing control parameter contours in violin performance. *Intelligent Data Analysis, 14*(5), 587-599.

Maezawa, A., Itoyama, K., Komatani, K., Ogata, T., Okuno, H. G., (2012). Automated violin fingering transcription through analysis of an audio recording. *Computer Music Journal, 36*(3), 57-72.

Mark, M.L., (1982). The evolution of music education philosophy from utilitarian to aesthetic. *Journal of Research in Music Education, 30*(1), 15-21.

McPherson, G. E., (1995). The assessment of musical performance: development and validation of five new measures. *Psychology of Music, 23*(2), 142-161.

McPherson, G. E., & O'Neill, S. A., (2010). Students' motivation to study music as compared to other school subjects: a comparison of eight countries. *Research Studies in Music Education, 32*(2), 101-137.

McPherson, G. E., and Thompson, W. F., (1998). Assessing music performance: issues and influences. *Research Studies in Music Education, 10*(1), 12-24.

McQuarrie, S. H., & Sherwin, R. G., (2013). Assessment in music education: relationships between classroom practice and professional publication topics. *Research and Issues in Music Education, 11*(1), 1-13.

Metcalf, C. D., Irvine, T. A., Sims, J. L., Wang, Y. L., Su, A. W. Y., Norris, D. O., Delogu, F., & Lee, R., (2014). Complex hand dexterity: a review of biomechanical methods for measuring musical performance. *Frontiers in Psychology, 5*, 414.

Meyer, J., & Land, R., (2003*). Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practising within the disciplines*: Edinburgh, University of Edinburgh.

Millican, S. P., (ed.) (2015). *Curriculum and assessment in preservice music teacher education*, Chicago: GIA.

Mills, J., (1991). Assessing musical performance musically. *Educational Studies, 7*(2), 173-181.

Mion, L., De Poli, G., (2008). Score-independent audio features for description of music expression. *IEEE Transactions on Audio, Speech, and Language Processing, 16*(2), 458–466.

Mishra, J., (2014). Factors related to sight-reading accuracy: a meta-analysis. *Journal of Research in Music Education*, *61*(4), 452-465.

Mitchell, D. C., Green, D.W., (1978). The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology*, *30*(4), 609-36.

Miyamoto, K. A., (2007). Musical characteristics of preschool-age students: a review of literature. *UPDATE: Applications of Research in Music Education, 26*(1), 26-40.

Moeslund, T. B., Hilton, A., Krüger, V., (2006). A survey of advances in vision-based human motion capture and analysis. *Computer vision and image understanding, 104*(2), 90-126.

Molina, E., Barbancho, I., Gómez, E., Barbancho, A., Tardón, L., , (2013, May). Fundamental frequency alignment vs. note-based melodic similarity for singing voice assessment. [Paper presentation] in *Proceedings* of the Institute of Electrical and Electronics Engineers International Conference. Acoustics, Speech and Signal Processing, Vancouver, Canada. (pp. 744-748).

Montgomery, D.C., (1992). *Introduction to linear regression analysis*. New York: Wiley.

Moog, H., (1976). The development of musical experience in children of pre-school age. *Psychology of Music, 4*(2), 38-45.

Moore, G., (2012). "Tristan chords and random scores": exploring undergraduate students' experiences of music in higher education through the lens of Bourdieu. *Music Education Research, 14*(1), 63-78.

Morgan, J., (1981). Sharpen your edge on choral competition. *Music Educators Journal, 67*(8), 44-47.

Morgan, K., (2011). Where is von Humboldt's university now?' research in higher education. *Daigaku Ronshu, 42*, 325-334

Muijs, D., (2004). *Doing quantitative research in education: with SPSS*. UK: SAGE Publications.

Musco, A. M., (2011). Beyond "line by line": strategies for performance and learning transfer. *Music Educators Journal, 38*(1), 59-67.

Nakano, T., Goto, M., Hiraga, Y., (2006, September). An automatic singing skill evaluation method for unknown melodies using pitch interval accuracy and vibrato features. [Paper presentation] in *Proceedings* of the International Conference on Spoken Language Processing Pittsburgh. PA, USA (pp. 1706–1709).

National Council for Curriculum and Assessment (NCCA) (1999). *Primary School Curriculum*, Dublin: NCCA

National Council for Curriculum and Assessment (NCCA) (2007). *Assessment in the primary school curriculum*, Dublin: NCCA

Neubäcker, P., (2011*). "What does a stone sound like?" a film portrait of Melodyne inventor Peter Neubäcker in conversation with Maximilian Schönher*. Available from https://www.youtube.com/watch?v=4pdSYkI86go. [accessed 21st April 2017].

Ng, K., Weyde, T., Larkin, O., Neubarth, K., Koerselman, T., Ong, B., (2007, November). The 3d augmented mirror: A multimodal interface for learning and teaching string instruments. [Paper presentation] in *Proceedings* of the 9th International Conference on Multimodal Interfaces, Nagoya, Aichi, Japan (pp. 339-345).

Nielsen, F. V., (2005). Didactology as a field of theory and research in music education. *Philosophy of Music Education Review, 13*(1), 5-19.

Nielsen, F. V., (2007) Music (and Arts) Education from the Point of View of Didaktik and Bildung, in (ed.) L. Bresler, *International Handbook of Research in Arts Education* (pp. 265-285). Dordrecht: Springer.

Nielsen, F. V., (2009). What is the significance of research for music education in practice? on relations between the practice of and the scientific approach to music education. *Philosophy of Music Education Review, 17*(1), 1-22.

Norris, C. E., (2013). Elementary children's tonal awareness as related to perception of tonal dissonance. *UPDATE: Applications of Research in Music Education, 31*(2), 63-69.

Oare, S., & Bernstorf, E., (2010). Literacy in the instrumental classroom. *Kodaly Envoy, 37*(1), 17-20.

Oliver, P., (2004) *Writing your thesis*. London: Sage.

Orman, E. K., Yarbrough, C., Neill, S., Whitaker, J. A., (2007). Time usage of middle and high school band directors in sight-reading adjudication. *UPDATE: Applications of Research in Music Education, 25*(2), 36-46.

Ortmann, O., (1926). On the melodic relativity of tones. *Psychological Monographs, 35*(1), 1-47.

Palkki, J., Albert, D. J., Hill, S. C., Shaw, R. D., (2016). 20 years of the MENC Biennial Conference. *Journal of Research in Music Education, 64*(1), 14-28.

Pardue, L. S., Harte, C., McPherson, A., (2015). A low- cost real-time tracking system for violin. *Journal of New Music Research, 44*(4), 305-323.

Parncutt, R., & McPherson, G., (ed.) (2002). *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning*. London: Oxford University Press.

Patterson, B., (1974). Musical dynamics. *Science America, 233*, 78-95.

Pellegrino, K., Conway, C. M., Russell, J. A., (2015). Assessment in performance-based secondary music classes. *Music Educators Journal, 102*(1), 48-55.

Phillips-Silver, J., Toiviainen, P., Gosselin, N., Peretz, I., (2013). Amusic does not mean unmusical: beat perception and synchronization ability despite pitch deafness. *Cognitive Neuropsychology, 30*(5), 311-331.

Prange, K., (2004) *Bildung:* a paradigm regained? *European Educational Research Journal*, *3*(2), 501-509.

Punch, K. F., (2006). *Developing effective research proposals*. London: Sage.

Quinlan, A.M., (2006) *A Complete Guide to Rubrics: Assessment Made Easy for Teachers*, Lanham: Rowman and Littlefield Education.

Ramella, M., Rabuffetti, M., Converti, R. M., Ferrarin, M., Fischer, S., Gardner, R. N., (1995). Quantitative analysis of ergonomical factors affecting violin player movement and posture. *Gait and Posture, 29*(1), 1-8.

Rasamimanana, N., (2012). Towards a conceptual framework for exploring and modelling expressive musical gestures. *Journal of New Music Research, 29*(3), 225-233.

Rcampus sight reading rubrics. (undated) Available from https://www.rcampus.com/rubricshowc.cfm?code=V98BX9&sp=yes& [accessed 24th February 2019].

Reed, R., (1990). Here comes the judge: solo string adjudication. *American String Teacher, 40*(3), 46-49.

Regenspurger, K., & Seidel, E. J., (2015). Prevention of musician-specific disorders of the musculoskeletal system. *Manuelle Medizin, 53*(1), 26-30.

Reifinger, J., (2006). Skill development in rhythm perception and performance: a review of literature. *UPDATE: Applications of Research in Music Education, 15*(1), 25-33.

Reifinger, J., (2009) An analysis of tonal patterns used for sight-singing instruction in second-grade general music class. *Journal of Research in Music Education, 57*(3), 203-216.

Remington Graphics (2018) DIY Kinect motion capture studio, Available from http://remingtongraphics.net/tutorials/diy-kinect-motion-capture-studio/, [accessed 23rd July 2018].

Repp, B. H., (2006). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic Bulletin & Review, 12*(6), 969-992.

Richmond, J., (2002). *Law research and music education*. New York: Oxford University Press.

Rickard, N. S., Bambrick, C.J., Gill, A., (2012). Absence of widespread psychosocial and cognitive effects of school-based music instruction in 10–13-year-old students. *International Journal of Music Education, 30*(1), 57-78.

Rosenberg, M., (1965). *Society and the adolescent self-image*. Middletown, CT: Wesleyan University Press.

Rousseau, J-J., (1762) *Emile, ou de l' Education*, Paris: Nicolas Bonaventure Duchesne.

Rousseau, J-J., (2002). *The Social Contract and The First and Second Discourses*, New Haven and London: Yale University Press.

Rowson, J., (2019). Bildung in the 21st Century -Why sustainable prosperity depends upon reimagining education *CUSP essay series on the Morality of Sustainable Prosperity, No 9,* Guildford: Centre for the Understanding of Sustainable Prosperity.

Russell, J., & Austin, J., (2010). Assessment practices of secondary music teachers. *Journal of Research in Music Education, 58*(1), 37-54.

Sadler, D. R., (2015). Backwards assessment explanations: implications for teaching and assessment practice. In D. Labler, et al., (eds) *Assessment in Music Education: from Policy to Practice* (pp. 9-19). Switzerland: Springer.

Sahlberg, P., (2015). *Finnish lessons*: *what can the world learn from educational change in Finland?* N.Y.: Teachers College Press.

Said, E., (2008). *Music at the limits*. London: Bloomsbury Publishing Plc.

Salaman, W., (1994). The Role of Graded Examinations in Music, *British Journal of Music Education, 11*(3), 209-221.

Scaife, N., (2015). *Bowed strings syllabus, 2016-2019*. London: ABRSM.

Schneider, J., Borner, D., Van Rosmalen, P., Specht, M., (2015). Augmenting the senses; a review on sensor-based learning support. *Sensors, 15*(12), 4097- 4133.

Schoonderwaldt, E., (2009a). Mechanics and acoustics of violin bowing: freedom, constraints and control in performance. Doctoral Thesis, Stockholm: KTH. Available from.http://www.divaportal.org/smash/get/diva2:133385/FULLTEXT01.pdf [accessed 3rd May 2019].

Schoonderwaldt, E., (2009b). The player and the bowed string: coordination of bowing parameters in violin and viola performance. *Journal of the Acoustical Society of America, 126*(5), 2709-2720.

Schoonderwaldt, E., (2010). On the use of skewness in violin bowing: should the bow be straight or not? *Acta Acustica united with Acustica, 96*(4), 593-602.

Schoonderwaldt, E., and Demoucron, M., (2009). Extraction of bowing parameters from violin performance combining motion capture and sensors. *Journal of the Acoustical Society of America, 126*(5), 2695-2708.

Seashore, C., E., (1915). The Measurement of Musical Talent *The Musical Quarterly*, *1*(1), 129-148.

Seashore, C. E., (1938). *Psychology of music*, New York: Dover.

Seashore, C. E., Lewis, L., Saetveit, J.G., (1960). *Seashore measures of musical talents*, New York: The Psychological Corporation.

Sergeant, D., (1973). Measurement of Pitch Discrimination. *Journal of Research in Music Education, 21*(1), 3-19.

Sergeant, D., & Boyle, J., (1980). Contextual influences on pitch judgement. *Psychology of Music, 8*(2), 3-15.

Sheridan, M. M., (2015). Teaching tips - general music. *Triad, 83*, 26.

Shuter-Dyson, R., (1968). *The psychology of musical ability*. London: Methuen.

Shuter-Dyson, R., & Gabriel, C., (1981). *The psychology of musical ability*. 2nd Edition. London: Methuen.

Sims, W. L., (1985). Young children's creative movements to music. *Contributions to Music Education, 1*(12), 42-50.

Slater, J., Strait, D. L., Skoe, E., O'Connell, S., Thompson, E., Kraus, N., (2014). Longitudinal effects of group music instruction on literacy skills in low-income children. *PLoS ONE, 9*(11), e113383.

Sloboda, J. A., (1984) Review of Wing and Seashore tests, In U. Frith (ed.) *Cognitive processes in spelling*. (pp. 231-248) London: Academic Press.

Sloboda, J. A., (1985). *The musical mind the cognitive psychology of music*. U.S.A.: Oxford University Press.

Sloboda, J. A., (2012) Patriotism and Nationalism in Music Education: A Review Essay, *International Journal of Education and the Arts, 13*(2), 1-6.

Small, A. M., (1937). An objective analysis of artistic violin performance. *University of Iowa studies in the psychology of music: Vol 4. Objective analysis of musical performance* (pp.172-231). Iowa City: University of Iowa.

Smith, K. C., Cuddy, L.L., Upitis, R., (1994). Figural and metric understanding of rhythm. *Psychology of Music, 22*(2), 117-135.

Smith, R., (2011). Beneath the skin: statistics, trust, and status. *Educational Theory, 61*(6), 633-645.

Steinbrink, C., Zimmer, K., Lachmann, T., Dirichs, M., Kammer, T., (2014). Development of rapid temporal processing and its impact on literacy skills in primary school children. *Child development, 85*(4), 1711-1726.

Streiner, D.L., (2005). Finding our way: an introduction to path analysis. *Research Methods in Psychiatry, 50*(2), 115-122.

Sun, J., Li, H., Ma, L., (2011). A music key detection method based on pitch class distribution theory. *International Journal of Knowledge Based Intelligent Engineering Systems, 15*(3), 165-175.

Swanwick, K., (1996). Teaching and assessing. *Newsletter of the Special Research Interest Group in Measurement and Evaluation (MENC), 18,* 6-9.

Tan, Y., Mc Pherson, G., Peretz, I., Berkovic, S., Wilson, S., (2014). The genetic basis of music ability. *Frontiers in Psychology. 5*(658), 1-18.

theEngineer, (2016). Sound used to map 3D-printing process, 3rd March, Available from http://www.theengineer.co.uk/sound-used-to-map-3d-printing-process/?cmpid=tenews_2087518)[accessed 20th Nov 2017].

Thompson, S., (2003). Evaluating evaluation: music performance assessment as a research tool. *Music Perception, 21*(1), 21-41.

Thomsen, L., (2012). Improve your violin or viola intonation in 5 simple steps. *Strings, 27*, 23.

Thomson, A., (2013). Following the Money – Review: The Associated Board of the Royal Schools of Music: A Social and Cultural History by David C.H. Wright, *The Musical Times, 154*(1923), 111-116.

Tomlinson, M. M., (2015). Transmodal redesign in music and literacy: Diverse multimodal classrooms. *Journal of Early Childhood Literacy, 15*(4), 533-567.

Trehub, S. E., Cohen, A. J., Thorpe, L. A., and Morrongiello, B. A., (1986). Development of the perception of musical relations: semitone and diatonic structure. Journal of Experimental Psychology: *Human Perception and Performance, 12*(3), 295-301.

Tzanetakis, G., and Cook, P., (2002). Musical genre classification of audio signals. *IEEE Transactions on Audio, Speech and Language Processing, 10*(5), 293-302.

Upitis, R., (1987). Children's understanding of rhythm: the relationship between development and music training. *Psychomusicology*: *A Journal of Research in Music Cognition, 7*(1), 41-60.

van der Linden, J., Schoonderwaldt, E., Bird, J., Johnson, R., (2011a). Musicjacket—combining motion capture and vibrotactile feedback to teach violin bowing. *IEEE Transactions on Instrumentation & Measurement, 60*(1), 104-113.

van der Linden, J., Johnston, R., Bird, J., Rogers, Y., Schoonderwaldt, E., (2011b, May). Buzzing to play: lessons learned from an in the wild study of real-time vibrotactile feedback. [Paper presentation] in *Proceedings* of the 29th International Conference on Human Factors in Computing Systems, Vancouver, BC, Canada (pp. 533-542).

van Manen, M., (1990). *Researching lived experience: Human science for an action sensitive pedagogy.* N.Y.: State University of New York Press.

Varkøy, Ø., (2015). Bildung: between cultural heritage and the unknown, instrumentalism and existence, In L. Bresler (Ed.), *The Routledge International Handbook of Arts and Education*. (pp. 19-29) London: Routledge.

Vurma, A., Raju, M., Kuuda, A., (2011). Does timbre affect pitch?: estimations by musicians and non-musicians. *Psychology of Music, 39*(3), 291.

Walford, G., (2001). *Doing qualitative educational research*. UK: Bloomsbury Academic.

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., Vuust, P., (2010). The musical ear test, a new reliable test for measuring musical competence. *Learning and Individual Differences, 20*(3), 188-196.

Waller, D., (2010). Language literacy and music literacy: a pedagogical asymmetry. *Philosophy of Music Education Review, 18*(1), 26-44.

Wang, J., Wang, S., Chen, W., Chang, K., Chen, H., (2012, August). Real-time pitch training system for violin learners. [Paper presentation] in *Proceedings* of the IEEE International Conference on Multimedia and Expo Workshops, Melbourne, Australia (pp. 163–168).

Waters, A. J., Underwood, G., (1998). Eye movements in a simple reading task. *Psychology of Music, 26*(1), 46-60.

Wesolowski, B. C., (2012). Understanding and developing rubrics for music performance assessment. *Music Educators Journal 98*(3), 36-42.

Wesolowski, B. C., (2015). Tracking student achievement in music performance: developing student learning objectives for growth model assessments. *Music Educators Journal, 102*(1), 39-47.

Wesolowski, B. C., (2016). Rater Precision in Music Performance Assessment: An Analysis of Rating Scale Structure Using the Multifaceted Rasch Partial Credit Model *Music Perception*, 33(5), 662-678

Wesolowski, B. C., (2019). Predicting Operational Rater-Type Classifications Using Rasch Measurement Theory and Random Forests: A Music Performance Assessment Perspective *Journal of Educational Measurement*, 56(3), 610-625.

Wesolowski, B. C., Wind, S.A. and Engelhard Jr, G., (2015). Rater fairness in music performance assessment: Evaluating model-data fit and differential rater functioning. *Musicae Scientiae, 19*(2), 147-170.

Whistler, H. S., and Thorpe, L. P., (1950). *Musical aptitude test*. Los Angeles: Calif. Test Bureau.

Whybrew, W. E., (1971). *Measurement and evaluation in music*. Iowa, USA: Wm.C. Brown Company Publishers.

Wiggins, D., (2007). Pre-k music and the emergent reader: promoting literacy in a music-enhanced environment. *Early Childhood Education Journal, 35*(1), 55-64.

Willems, E., (2012). *Psychological foundations of musical education* 4th Edition. Translation, Psychologiques de l'éducation musicale by Jerry-Louis Jaccard. Provo, Utah: Creative Works.

Wind, S. A., and Engelhard, G., (2015). Exploring rating quality in rater-mediated assessments using mokken scale analysis. *Educational and Psychological Measurement, 76*(4), 685-706.

Wing, H. D., (1948). Tests of musical ability and appreciation: an investigation into the measurement, distribution, and development of musical capacity. *Volume 27 of British Journal of Psychology, Monograph Supplements,* Cambridge: Cambridge University Press

Wing, H. D., (1962). *Wing standardized tests of musical intelligence,* Windsor: National Foundation for Educational Research.

Wright, D., (2013). *The Associated Board of the Royal Schools of Music: A Social and Cultural History*, Woodbridge: The Boydell Press.

Wright, S., (1934). The method of path coefficients. *The Annals of Mathematical Statistics 5*(3), 161-215.

Wrigley, W., (2011). Ecological development and validation of a music performance rating scale for five instrument families. *Psychology of Music, 41*(9), 97-118.

Wu, C., Laguna, C., Pati, A., Vidwans, A., and Lerch, A., (2016, July). Towards the objective assessment of music performances. [Paper presentation] in *Proceedings* of the 14th International Conference on Music Perception and Cognition (ICMPC), San Francisco, USA (pp. 99-102).

Yea, C., Roebel, A., Rodet, X., (2010). Multiple fundamental frequency estimation and polyphony interference of polyphonic music signals. *IEEE Transactions on Audio, Speech, and Language Processing, 18*(2), 1116-1126.

Zdzinski, S. F., & Barnes, G. V., (2002). Development and validation of a string performance rating scale. *Journal of Research in Music Education, 50*(3), 245-255.

Zelić, T., (2018). Bildung and the historical and genealogical critique of contemporary culture: Wilhelm von Humboldt's neo-humanistic theory of Bildung and Nietzsche's critique of neo-humanistic ideas in classical philology and education, *Educational Philosophy and Theory*, *50*(6-7), 662-671.

Zenatti, A., (1969). Le developpement genetique de la perception musicale. *(Monographies Francaises de Psychology.,* no. 17) Paris: Centre National de Recherche Scientifique.

Zhang, B., Wang, Y., (2009). *Automatic music transcription using audio-visual fusion for violin practice in home environment*. Singapore: National University of Singapore.

Appendix 1

Research Study
Participant Information and Consent Form

I am a music specialist with thirty years teaching and performance experience. I have a Master of Arts degree in string performance and am currently a doctoral research student at the University of Lincoln. I am conducting an observational study into musical perception and attainment. It is hoped that my research will be of interest to pre-service generalist and specialist primary school teachers in teacher training programmes in the future. It is also expected that the findings will influence policy makers and curriculum design.

I will be collecting data on musical attainment markers of fourth and fifth class students following four years of classroom learning. The purpose of the study is to ascertain the efficacy of group lessons in a classroom context, with regard to predetermined learning goals in string playing. It is also an objective of the study to predict outcomes at future grade levels, and help to guide strategies into the future.

Students will be observed objectively for criteria that fall into five separate categories of learning. They include: pitch (detecting something being out of tune), rhythm (the ability to play in time with others), posture (relating to a sustainable ergonomic approach to holding the violin, its tone production and a bowing strategy), intonation (the ability to play in tune), and reading (the ability to interpret written indications of notes, rhythmic detail and bowing content).

This is not a test of the student, nor does it reflect in any way on the child's ability to perform music. Rather, it is an evaluation of teaching practice and learning milestones in a classroom context. Musicianship, virtuosity and competitive considerations are outside the scope of this study, as focus is explicitly placed on a music learning process, rather than music learnt.

Students, having been advised of the tasks in advance, will be asked to: listen and respond to short extracts, perform a simple scale, read a suitably easy passage of music, play a piece of their own choice and, with their musical partner, perform a piece of music which features on the child's reading list. Students will present in pairs for five five-minute sessions between September and December 2015. Music classes will not be affected, as the observations are separate entities from music classes.

Confidentiality is assured, with student identity being coded and anonymity established. The data will be collated with grade music examinations taken in April 2016. Data will be stored securely on a password-protected hard drive and an anonymous copy will be retained for five years. Midi technology using Sibelius software will be incorporated in the pitch, reading and rhythm observations, and motion capture technology will be used to analyse posture and bow angle accuracy. Openness and transparency is built into the design of the research methodology and each stage can be followed on request.

The school currently provides string tuition to over one hundred and fifty students, with one hundred and thirty students taking home specially-sized violins to practise on. It is hoped that the findings of this study will help to build a case for music education programmes at national level and address the continuum issue in music education.

Patrick Early August 2015

I give consent to my child participating in the music study □
I do not give consent to my child participating in the music study □

Signature of Primary Carer ________________________

Signature of Child ________________________

Appendix 2

Classroom Tasks Learning Structure

Junior Infants Tasks

1. Movement: The children respond to music played by dancing fast or slow, while imitating high, low, loud and quiet visual responses through movement. This is followed by 'musical chairs', 'musical statues', 'duck duck goose' and other games which stimulate attention and reaction to audio stimulus.
2. Rhythm: The 'Quack Rhythm' booklet and whiteboard projection charts (see Book 1), and the accompanying 'Quack Mat' movement in rhythm exercise, prepare the children for rhythmic concepts before contact with the instrument, as a precursor to music notation learning.
3. Posture: Beginnings relating to how the bow is held, the 'Woof Woof Stick' and 'Spider Walk' are practices introduced to orientate the students to a correct bow hold. Holding a piece of aero board under the chin whilst at the same time clapping the hands, instils confidence in holding the instrument without full support of the left hand at a later stage. These important overtures lay the foundation for the child being given their own 1/8 size instrument in the following year, at age five.

Senior Infants Tasks

The students are introduced to their own sized instrument. Bowing open strings in response to open strings heard, starting with the outer strings (green and pink). Plucking open strings in response to open strings heard. Reading notation colours for the open strings: G green, D red, A blue and E pink (see Book 2). Issues associated with holding the instrument and instrument care are also discussed. Reading from letter names is not recommended, as it postpones a music literacy platform at a crucial learning stage. The intermediary language of letters to represent sounds becomes an obstacle to music literacy when it precedes a reflex system and, in many cases, postpones or eclipses literacy altogether.

First, Second and Third Class Tasks

The fingers of the left hand are introduced at this stage, building on the open strings note pictures from the previous class. Repetition of short, four-note motifs is a feature of this stage of learning. The workout with the fingers (see Book 3) orientates the student to adopt the same reflex response to note pictures, but using fingers with appropriate colours for a given string. Finger patterns which place the 2nd finger beside the 3rd finger are adopted throughout, despite the contradictory musical context on the E string (G sharp on the E string being acceptable despite no key signature being present. Key signatures become more relevant from Second and Third Class onwards). These grammatically incorrect steps are taken in First Class to avoid cognitive overload. The short motifs revisit rhythmic features encountered in Book 1.

Repetition of longer two-bar motifs in Second Class build simple tune structures from memory. Both through responding to tunes played, and the coloured notation which represents the notes needed to play them, the student begins to play tunes which are already familiar (nursery rhymes and folk melodies), introduced in Book 4. At this stage, special attention should be given to individuals who may be struggling to hold the instrument correctly or place their fingers on the strings. Attention to bow hold and posture development is recommended to avoid problems later. Preparation and performance of these tunes is encouraged to build confidence and self-esteem in working from memory and notation.

Children at Third Class level returning to class in September, having been on the programme, should be encouraged to participate in a Christmas concert in early December. The reading material for this level is prepared with this in mind, in Book 5. While most students will want to play the familiar tunes associated with the festive season, an easy accompaniment part is included for each piece, to encourage students who would like to participate but feel unable to play the melody. Children from Second Class (and, in some cases, First Class) can perform these parts if they wish. An interactive whiteboard that can follow the music has been developed to facilitate the association between note picture and accompaniment track in concert preparations, as in all of the learning stages. Approaches to writing and sharing students' musical ideas are also explored in this class, with compilations of ideas being reworked to feature in performance. See Books 5.1 and 5.2.

Fourth, Fifth and Sixth Class Tasks

Children at this level are striving to improve tone production and intonation. Suitable material for this level must strike a balance between process and product. Many pieces have been explored and rejected because of class dynamics and differentiation issues (working at different speeds of development for different ability students). The resulting Book 6 strives to strike a balance between musicality and difficulty. Moves away from colour to black and white notation are discussed, with tasks to learn note names and fingers associated with them at this stage being a requirement for progress to be made to the next level. Material for Prep Test (pre-grade) standard is given to the students in anticipation of assessment following six years of classwork study.

Students on the programme move from the colour format of the earlier books to reading continuously in black and white in Fifth Class, as this will be needed in Grade I preparations. Extracts from the syllabus have been prepared for class instruction and two-part performance tasks are introduced to encourage students to develop a conversational approach to their music-making, as featured in Book 7. Students are coached in performance routines and rehearse with accompaniment tracks, before working with a live accompanist, prior to the exam in this year.

The final class cohort begin preparation for a Grade II practical examination. In addition to this, new material at Sixth Class includes four-part writing, with material for violins and cello performance prepared in Book 8. Some of the pieces have been taken from the syllabus as an added incentive to students to perform ensemble, as with piano, for their Grade II performance. Music theory is also introduced for the first time, with discussions about key and time signatures, finger patterns relating to the different keys and aural training.

Appendix 3

Raw Data

Name	ID	Pitch (+)			Intonation (-)				Rhythm (+)			Instrument Angle (-)			
		Raw Score	Optimum	% Score	Raw Score	Units of Excellence	Optimum	% Score	Raw Score	Optimum	% Score	Raw Score	U.E.	Optimum	% Score
Aine	1	7	10	70	328	1172	1500	78	35	40	88	20	40	0 of 60	67
Megan	3	9	10	90	220	1280	1500	85	38	40	95	10	50	0 of 60	83
Fionn	4	3	10	30	109	1391	1500	93	39	40	98	10	50	0 of 60	83
Claire	5	6	10	60	302	1198	1500	80	39	40	98	20	40	0 of 60	67
James	10	5	10	50	160	1340	1500	89	5	40	13	40	20	0 of 60	33
Emma K	11	7	10	70	263	1237	1500	82	24	40	60	10	50	0 of 60	83
Eillie	12	5	10	50	192	1308	1500	87	21	40	53	20	40	0 of 60	67
Oscar B	14	4	10	40	197	1303	1500	87	36	40	90	10	50	0 of 60	83
Finbar	15	7	10	70	83	1417	1500	94	29	40	73	10	50	0 of 60	83
Finnaula	16	8	10	80	46	1454	1500	97	32	40	80	5	55	0 of 60	92
Cian	17	9	10	90	160	1340	1500	89	30	40	75	20	40	0 of 60	67
Ruan	18	8	10	80	147	1353	1500	90	27	40	68	10	50	0 of 60	83
Ciaran	21	8	10	80	57	1443	1500	96	31	40	78	20	40	0 of 60	67
Aibfinn	22	6	10	60	79	1421	1500	95	30	40	75	10	50	0 of 60	83
Seosamh	23	8	10	80	311	1189	1500	79	33	40	83	20	40	0 of 60	67
Gretta	27	4	10	40	262	1238	1500	83	22	40	55	30	30	0 of 60	50
Geraldine	29	4	10	40	231	1269	1500	85	24	40	60	20	40	0 of 60	67
Aisling	30	7	10	70	237	1263	1500	84	34	40	85	10	50	0 of 60	83
Niamh	31	5	10	50	315	1185	1500	79	30	40	75	15	45	0 of 60	75
Cieran 2	32	4	10	40	178	1322	1500	88	23	40	58	20	40	0 of 60	67
Micheal	33	5	10	50	236	1264	1500	84	30	40	75	15	45	0 of 60	75
Roselle	34	7	10	70	52	1448	1500	97	36	40	90	5	55	0 of 60	92
Maria	35	6	10	60	189	1311	1500	87	31	40	78	5	55	0 of 60	92
Aoife O	36	9	10	90	80	1420	1500	95	38	40	95	5	55	0 of 60	92
Sophia S	37	6	10	60	121	1379	1500	92	33	40	83	30	30	0 of 60	50
Sara S	38	8	10	80	151	1349	1500	90	32	40	80	45	15	0 of 60	25
Sara B	39	7	10	70	79	1421	1500	95	34	40	85	10	50	0 of 60	83
Christian	40	8	10	80	41	1459	1500	97	36	40	90	5	55	0 of 60	92
Marese	41	8	10	80	37	1463	1500	98	38	40	95	10	50	0 of 60	83
Grainne	42	7	10	70	79	1421	1500	95	32	40	80	10	50	0 of 60	83
Lea Mei	43	9	10	90	40	1460	1500	97	36	40	90	5	55	0 of 60	92
Sophia R	44	7	10	70	68	1432	1500	95	35	40	88	5	55	0 of 60	92
Honour	45	7	10	70	159	1341	1500	89	29	40	73	10	50	0 of 60	83
Roisin	46	3	10	30	262	1238	1500	83	28	40	70	10	50	0 of 60	83
Ellen	47	6	10	60	95	1405	1500	94	30	40	75	10	50	0 of 60	83
Siofra	48	6	10	60	265	1235	1500	82	33	40	83	17	43	0 of 60	72
Molly	49	7	10	70	151	1349	1500	90	33	40	83	16	44	0 of 60	73

Nut Elbow (-)					Othogonal (-)						Mid Elbow (-)					Mid Orthogonal (-)						Point Elbow (+)				
aw Score	Optimum	Offsett	U.E.	% Score	Raw Score	Optimum	Offset	Normalise	U.E.	% Score	Raw Score	Optimum	Offset	U.E.	% Score	Raw Score	Optimum	Offset	Normalise	U.E.	% Score	Raw Score	Optimum	Offset	U.E.	% Score
97	60	37	23	38	106	90	16	16	74	82	103	90	13	77	86	81	90	9	9	81	90	89	160	71	89	56
98		38	22	37	92	90	2	2	88	98	135		45	45	50	81		9	9	81	90	146		14	146	91
88		28	32	53	89	90	-1	1	89	99	112		22	68	76	87		3	3	87	97	115		45	115	72
90		30	30	50	100	90	10	10	80	89	107		17	73	81	83		7	7	83	92	121		39	121	76
100		40	20	33	108	90	18	18	72	80	120		30	60	67	78		12	12	78	87	96		64	96	60
81		21	39	65	84	90	-6	6	84	93	95		5	85	94	85		5	5	85	94	135		25	135	84
95		35	25	42	77	90	-13	13	77	86	98		8	82	91	71		19	19	71	79	134		26	134	84
79		19	41	68	75	90	-15	15	75	83	98		8	82	91	69		21	21	69	77	134		26	134	84
80		20	40	66	95	90	5	5	85	94	115		25	65	72	82		8	8	82	91	144		16	144	90
65		5	55	91	90	90	0	0	90	100	90		0	90	100	89		1	1	89	99	120		40	120	75
133		73	-13	-22	90	90	0	0	90	100	111		21	69	77	78		12	12	78	87	128		32	128	80
93		33	27	45	92	90	2	2	88	98	94		4	86	96	98		-8	8	82	91	110		50	110	69
83		23	37	61	92	90	2	2	88	98	119		29	61	68	78		12	12	78	87	130		30	130	81
72		12	48	80	91	90	1	1	89	99	121		31	59	66	83		7	7	83	92	138		22	138	86
101		41	19	32	91	90	1	1	89	99	135		45	45	50	81		9	9	81	90	148		12	148	93
106		46	14	23	68	90	-22	22	68	76	107		17	73	81	67		23	23	67	74	123		37	123	77
98		38	22	37	93	90	3	3	87	97	102		12	78	87	100		-10	10	80	89	98		62	98	61
70		10	50	83	91	90	1	1	89	99	94		4	86	96	97		-7	7	83	92	120		40	120	75
103		43	17	28	94	90	4	4	86	96	98		8	82	91	95		-5	5	85	94	111		49	111	69
98		38	22	37	97	90	7	7	83	92	97		7	83	92	98		-8	8	82	91	108		52	108	68
83		23	37	61	93	90	3	3	87	97	93		3	87	97	95		-5	5	85	94	117		43	117	73
68		8	52	86	90	90	0	0	90	100	91		1	89	99	90		0	0	90	100	125		35	125	78
72		12	48	80	91	90	1	1	89	99	92		2	88	98	93		-3	3	87	97	121		39	121	76
91		31	29	48	90	90	0	0	90	100	90		0	90	100	89		1	1	89	99	128		32	128	80
95		35	25	42	100	90	10	10	80	89	102		12	78	87	79		11	11	79	88	100		60	100	63
104		44	16	27	98	90	8	8	82	91	99		9	81	90	86		4	4	86	96	98		62	98	61
75		15	45	75	92	90	2	2	88	98	94		4	86	96	88		2	2	88	98	123		37	123	77
80		20	40	66	91	90	1	1	89	99	92		2	88	98	97		-7	7	83	92	119		41	119	74
83		23	37	61	93	90	3	3	87	97	95		5	85	94	86		4	4	86	96	107		53	107	67
79		19	41	68	92	90	2	2	88	98	93		3	87	97	87		3	3	87	97	120		40	120	75
68		8	52	86	91	90	1	1	89	99	91		1	89	99	92		-2	2	88	98	130		30	130	81
72		12	48	80	91	90	1	1	89	99	93		3	87	97	96		-6	6	84	93	128		32	128	80
84		24	36	60	94	90	4	4	86	96	98		8	82	91	98		-8	8	82	91	113		47	113	71
81		21	39	65	98	90	8	8	82	91	97		7	83	92	101		-11	11	79	88	107		53	107	67
93		33	27	45	94	90	4	4	86	96	96		6	84	93	95		-5	5	85	94	116		44	116	73
104		44	16	27	75	90	-15	15	75	83	121		31	59	66	71		19	19	71	79	100		60	100	63
70		10	50	83	93	90	3	3	87	97	95		5	85	94	95		-5	5	85	94	108		52	108	68

Point Orthogonal (-)						Trajectory% (+) (bow totals)			Distance (+)				Guidance (-)				Notes (-)				Duration (-)			
Raw Score	Optimum	Offset	Normalised	U.E.	% Score	Raw Score	Optimum	% Score	Raw Score	Bow Size	Optimum	% Score	Raw Score	U.E.	Optimum	% Score	Raw Score	U.E.	Optimum	% Score	Raw Score	U.E.	Optimum	% Score
63	90	-27	27	63	70	422	600	70	772	62	930	83	5	15	20	75	3	17	20	85	4	16	20	80
96		6	6	84	93	459		76	535	62	930	57	4	16		80	0	20		100	1	19		95
105		15	15	75	83	479		80	919	62	930	98	0	20		100	0	20		100	0	20		100
120		30	30	60	67	454		76	535	62	930	57	1	19		95	1	19		95	1	19		95
124		34	34	56	62	389		65	852	62	930	91	6	14		70	5	15		75	4	16		80
99		9	9	81	90	521		87	611	62	930	65	0	20		100	1	19		95	1	19		95
104		14	14	76	84	465		78	557	62	930	60	2	18		90	2	18		90	1	19		95
123		33	33	57	63	466		78	785	62	930	84	0	20		100	0	20		100	0	20		100
122		32	32	58	64	479		80	861	62	930	92	1	19		95	0	20		100	0	20		100
93		3	3	87	97	562		94	717	62	930	77	0	20		100	0	20		100	0	20		100
111		21	21	69	77	398		66	716	62	930	77	2	18		90	0	20		100	1	19		95
98		8	8	82	91	489		82	942	64	960	98	0	20		100	0	20		100	0	20		100
111		21	21	69	77	472		79	795	62	930	85	0	20		100	0	20		100	0	20		100
116		26	26	64	71	494		82	810	62	930	87	0	20		100	0	20		100	0	20		100
121		31	31	59	66	428		71	959	68	1020	94	0	20		100	0	20		100	0	20		100
129		39	39	51	57	388		65	520	62	930	56	7	13		65	4	16		80	5	15		75
97		7	7	83	92	462		77	720	62	930	77	3	17		85	1	19		95	2	18		90
96		6	6	84	93	538		90	800	62	930	86	0	20		100	0	20		100	0	20		100
103		13	13	77	86	464		77	698	62	930	75	1	19		95	1	19		95	1	19		95
107		17	17	73	81	461		77	654	68	1020	64	2	18		90	1	19		95	2	18		90
106		16	16	74	82	505		84	730	62	930	78	1	19		95	0	20		100	0	20		100
91		1	1	89	99	562		94	800	62	930	86	0	20		100	0	20		100	0	20		100
94		4	4	86	96	544		91	765	62	930	82	0	20		100	0	20		100	1	19		95
92		2	2	88	98	525		87	800	68	1020	78	0	20		100	0	20		100	0	20		100
102		12	12	78	87	454		76	645	62	930	69	2	18		90	1	19		95	3	17		85
112		22	22	68	76	440		73	586	62	930	63	1	19		95	2	18		90	3	17		85
96		6	6	84	93	536		89	796	62	930	85	0	20		100	0	20		100	1	19		95
93		3	3	87	97	526		88	801	68	1020	78	0	20		100	0	20		100	0	20		100
99		9	9	81	90	505		84	821	68	1020	80	1	19		95	0	20		100	0	20		100
95		5	5	85	94	529		88	798	62	930	85	0	20		100	0	20		100	1	19		95
91		1	1	89	99	562		94	807	62	930	86	0	20		100	0	20		100	0	20		100
97		7	7	83	92	541		90	869	62	930	93	0	20		100	0	20		100	0	20		100
99		9	9	81	90	498		83	685	68	1020	67	2	18		90	1	19		95	2	18		90
93		3	3	87	97	499		83	543	64	960	56	3	17		85	1	19		95	3	17		85
95		5	5	85	94	495		83	748	62	930	80	1	19		95	2	18		90	1	19		95
129		39	39	51	57	373		62	708	62	930	76	3	17		85	3	17		85	2	18		90
112		22	22	68	76	512		85	910	62	930	97	1	19		95	0	20		100	1	19		95

Bowings (-)				Reading (+)	Grade (+)			Variable (+
Raw Score	Positives	Optimum	% Score	Total	Raw Score	Optimum	% Score	% Score
3	17	20	85	83	102	150	68	77
2	18		90	95	120		80	82
0	20		100	100	114		76	84
1	19		95	95	115		77	77
7	13		65	73	113		75	70
0	20		100	97	133		89	84
2	18		90	92	126		84	76
0	20		100	100	123		82	82
0	20		100	100	118		79	88
0	20		100	100	131		87	91
1	19		95	97	112		75	83
0	20		100	100	127		85	90
0	20		100	100	131		87	87
0	20		100	100	123		82	87
1	19		95	98	126		84	85
6	14		70	75	96		64	65
4	16		80	88	109		73	76
0	20		100	100	121		81	88
1	19		95	95	110		73	79
2	18		90	92	124		83	76
1	19		95	98	102		68	82
0	20		100	100	121		81	91
0	20		100	98	125		83	87
0	20		100	100	120		80	91
4	16		80	87	102		68	76
5	15		75	83	91		61	75
2	18		90	95	110		73	88
1	19		95	98	113		75	90
2	18		90	97	127		85	88
3	17		85	93	124		83	88
0	20		100	100	132		88	93
1	19		95	98	126		84	91
2	18		90	92	122		81	83
2	18		90	90	106		71	75
1	19		95	93	121		81	84
6	14		70	82	115		77	76
3	17		85	93	127		85	86

Calculations of variables and their conversion to percentage scores

Intonation
Zero is excellence
1500 is the worst
328 scored errors
therefore student got
1500-328 correct = 1172
1172 out of 1500 = (1172 x 100) / 1500 = 78.133
220 errors therefore score is 1500-220=1280
1280 correct out of 1500 = 85

Bow Angles
60 degrees = 100%
therefore 1 degree is 100/60 = 1.6666 recurring
if we have 97 degrees, this is an error of 97 – 60 =37 degrees
37 x 1.6666 = 61.64 we are measuring error, not excellence

60 degrees is perfect, but we need to express the level of perfection as a % out of 100.
If we have 97 degrees we are 37 degrees from perfection (i.e. 37 degrees from 60)
therefore, the extent of our perfection is the perfect score (60) less the degree of imperfection (37) =23
23 x 1.66 = 38.318 we are measuring the extent of excellence

Example
If we have an observation of 23 degrees, this is an error of 60-23 = 37
37 x 1.666 = 61.64
if we have an observation of 100 degrees, then we are 40 degrees from perfection so the score is 60-40= 20
20 x 1.666=33.32
if we have 87, this is an error of 87-60=27 degrees
27 x 1.666 = 44.98

following the above for orthogonal angles
90 degrees is 100%
therefore 1 degree is 100/90 =1.111
160degrees =100%
therefore 1 degree is 100/160 =.625

Bow Distance
½ Size
62cm x15=930cm
Excellence is 930cm normalised at 100%
100/930=.107 = 1 degree of perfection
observed =772
Normalised is 772 x.107 =82.6%

¾ Size
68.7cm x 15=1030cm
Excellence 1030cm normalised at 100%
100/1030 = .097 = 1 degree of perfection
Observed = 538
Normalised value is 538 x .097 = 52.18%

Guidance
Ideal = 0 over 20 elements
20 elements without errors
5 errors
15 notes with no errors
15 out of 20 is 75%

Reading
20 is excellence i.e. 20 elements without errors
Student made 3 errors, hence student played 17 notes correctly
17/20 =85

Instrument calculation
Ideal is 0 divergence over 60 degrees
1 degree = 100/60 = 1.666%
if we have angle error of 20 degrees off
that means that 40 degrees of possible 60 is correct
therefore 40 x 1.666 = 66.64%.