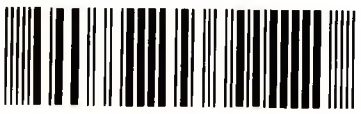
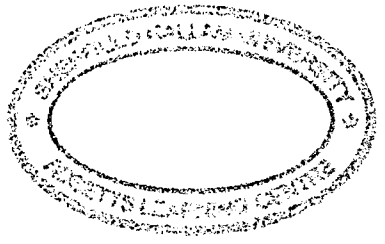


'Knowing Yourself Through Others': Peer Assessment in Popular Music Group Work

Mark Pulman

A thesis submitted in partial fulfilment of the requirements of
Sheffield Hallam University
For the degree of Doctor of Philosophy

March 2008



Abstract

This enquiry investigates the experiences and responses to peer assessment of group work involving cohorts of undergraduate popular music students over a five-year period. Working within the context of band rehearsing and performing, the enquiry focuses on how intra-peer assessment impact on students' personal attributes and their learning.

The literature review presents an overview of peer learning, group work, peer assessment processes, and a survey of the peer assessment literature on music in Higher Education reveals a lack of research into popular music group work.

An action research design was established to study developing peer assessment activities of group work involving nineteen rehearsing and performing cycles. This allowed interventions and refinements to be made from cycle to cycle from which qualitative interview data and quantitative peer assessment data were collected.

The analysis and interpretation of this data explain the key themes that arose from the students' experiences of peer assessment in the action research. These include the development of awareness and knowledge about their personal attributes. Confidence, feedback and a moral dimension, often involving honesty and trust, were of particular significance.

A new process model of intra-peer assessment is proposed. It offers a sequence of graduated stages of personal attribute usage, which create experiences over a period of time, that support students' learning about themselves and about others through intra-peer assessment activities. The key activity, which also gives the model its particular distinctiveness, involves bands decide for each of their members appropriate personal attributes to be used as criteria for intra-peer assessment.

The enquiry emphasises the importance of providing experiential and interactional contexts for intra-peer assessment, as important learning opportunities arise from such settings. This study provides a social constructivist explanation for the development of students' personal attributes and the building of trust and honesty in the rehearsing and performing cycles.

Acknowledgements

There are four individuals who I would like to thank especially for their involvement in this study.

Anthony Rosie (Professor of Social Science Education), for his guidance and supervision of the middle and final stages of this thesis; Professor Peter Hartley (now Professor of Education Development at the University of Bradford) for his initial supervisory input and group work ideas; Sue Drew, Principal Lecturer in Education, for her interest and encouragement throughout this enquiry. Finally, I would like to thank Lorna for her support and patience over the years.

Last, but not least, I am grateful to the considerable numbers of popular music students who participated in this study and made it possible.

Contents

Introduction 1

1. Literature Review 6

1.1 Introduction 7

1.2 Peer assessment and learning 10

1.3 Group work and peer assessment 17

1.4 The process of peer assessment 33

1.5 Peer assessment literature in the HE music curriculum 51

1.6 Epilogue: the claims and the pains of peer assessment 63

2. Methodological Considerations 69

2.1 History and context 73

2.2 Considering paradigmatic assumptions 75

2.3 Considering peer assessment: exploration, practice, issue and questions 80

2.4 Rehearsing and performing cycles 83

2.5 Considering mixed methods 84

2.6 Considering an action research approach 86

2.7 Role and ethnographic considerations 90

2.8 Consideration of values 93

2.9 Supporting learning 95

2.10 The curriculum area 95

2.11 Summary 96

3. Methods of Enquiry 97

3.1 Course, module details and students 99

3.2 Commencing action research 103

3.3 Developing the intra-peer assessment enquiry 106

3.4 Developing the inter-peer assessment enquiry 113

3.5 Developing the process features enquiry 116

3.6 Data collection and methods of analysis 116

3.7 Standards of quality and verification 131

4. The rehearsing and performing cycles 134

4.1 Illustrations of the cycles 136

4.2 Rehearsing and Performing Cycles: 2000-2001 138

4.3 Rehearsing and Performing Cycles: 2001-2002 141

4.4 Rehearsing and Performing Cycles: 2002-2003 144

4.5 Rehearsing and Performing Cycles: 2003-2004 148

4.6 Rehearsing and Performing Cycles: 2004-2005 152

4.7 Summary and analytical overview 155

5. Qualitative and Quantitative Data Analysis 159

5.1. Interview data analysis 161

5.2. Peer assessment data analysis 178

5.3. Relationships between the data analyses 195

6. 'Knowing yourself through others': a new process model for intrapeer assessment 201

6.1 Presenting a new process model of intrapeer assessment 203

6.2 Stage one 204

6.3 Stage two 206

6.4 Stage three 208

6.5 Stage four 210

6.6 Using personal attributes as assessment criteria: developing awareness and supporting learning 213

6.7 Moral dimensions of the process model 215

6.8 The process model and band determined personal weaknesses 219

6.9 Strengths and weaknesses of the model 223

7. The research in context 227

7.1 Examining the research through different perspectives 228

7.2 Implications of the action research for tutors 243

7.3 The curriculum context 245

7.4 The methodological context 248

7.5 The personal context 249

7.6 Summary of the contribution to original knowledge 253

8. Bibliography 254

9. Appendix 270

Tables and Figures

1. Literature Review

Table 1.1: *The Claims of Peer Assessment.* 64

Table 1.2: *The Pains of Peer Assessment.* 66

2. Methodological Considerations

Figure 2.1: *Emerging methodological considerations and assumptions and their interrelationships.* 71

Figure 2.2: *Synthesis of the elements of the enquiry.* 88

3. Methods of Enquiry

Table 3.1: *Summary of the courses and modules, research focus and student cohorts.* 103

Table 3.2: *Group rehearsing and performing cycles.* 104

Figure 3.1: *Intra-peer assessment form based on a general contribution to rehearsing criterion.* 107

Figure 3.2: *Semantic differential rating scale based upon bi-polar personal attributes.* 110

Figure 3.3: *Example of an intra-peer assessment form using group attributes and individual attributes.* 112

Figure 3.4: *Summary of the 'personal attribute' terms used in this enquiry.* 113

Table 3.3: *Summary of methods of enquiry, quantitative and qualitative data collection.* 117

4. The rehearsing and performing cycles

Figure 4.1: *The rehearsing and performing cycles in the action research.* 137

Figure 4.2: *Development of the rehearsing and performing cycles: 2000-2001.* 140

Figure 4.3: *Development of the rehearsing and performing cycles: 2001-2002.* 143

Figure 4.4: *Development of the rehearsing and performing cycles: 2002-2003.* 147

Figure 4.5: *Development of the rehearsing and performing cycles: 2003-2004.* 151

Figure 4.6: *Development of the rehearsing and performing cycles: 2004-2005.* 154

Table 4.1: *Model showing the relationship between the types of attributes employed in the rehearsing and performing cycles and the increasing action research focus.* 156

Table 4.2: *Band determined personal weaknesses and their impact on the increasing focus of the action research.* 157

5. Qualitative and Quantitative Data Analysis

Table 5.1: *Interviews with Performance Management students 2002-2005.* 161

Table 5.2: *Overview of the categories produced from the interview data analysis.* 162

Table 5.3: *Axial coding: eight core categories.* 163

Table 5.4: *Axial coding: core categories and sub-categories.* 164

Figure 5.1: *Selective coding foci of interview data.* 170

Figure 5.2: *The focus on self knowledge.* 172

Figure 5.3: *The focus on feedback.* 174

Figure 5.4: *The focus on confidence.* 175

Figure 5.5: *The focus on honesty.* 176

Table 5.5: *Overview of the analytical procedures.* 178

Table 5.6: *Peer assessment data 2000-2005: Propositional claims.* 183

Figure 5.7: *Interpretation of the quantitative data: group attributes.* 192

Figure 5.8: *Interpretation of the quantitative data: personal weaknesses.* 192

Figure 5.9: *Interpretation of the quantitative data: free riders.* 193

Figure 5.10: *Interpretation of the quantitative data: shooting stars.* 193

6. 'Knowing yourself as others know you': a new process model for intra-peer assessment

Figure 6.1: *Stages of the intra-peer assessment process model.* 204

Figure 6.2: *Stage Two of the intra-peer assessment process model.* 206

Figure 6.3: *Stage Three of the intra-peer assessment process model.* 209

Figure 6.4: *Stage Four of the intra-peer assessment process model.* 210

Figure 6.5: *Overview of the three stages of the intra-peer assessment process model.* 212

Figure 6.6: *Personal attributes arising from the qualitative and quantitative data* analyses *and their possible relationship to the intra-peer assessment process model.* 214

Table 6.1: *Summary of the relationship between trust, honesty and the process model.* 216

Figure 6.7: *Focus on band determined personal weaknesses arising from the enquiry.* 222

7. The research in context

Figure 7.1: *A symbolic interactionist perspective of the enquiry.* 234

Figure 7.2: *Epistemology of rehearsing interaction, choices and decisions.* 236

Figure 7.3: *Interactional opportunities for knowledge creation.* 237

Figure 7.4: *Meanings of honesty and trust in the 'dynamic relationship' involving students, tutor and the intra-peer assessment process model.* 239

Figure 7.5: *Possible syntheses of theoretical perspectives for the enquiry.* 240

Table 7.1: *Suggested epistemology of the enquiry relating to the 'active engagement'* 243

Table 7.2: *Types of personal attributes, focus of learning and goal.* 247

Introduction

Education is often characterised by its propensity for change. Curricular content, pedagogy and assessment, for example, seem to be popular and frequent targets for change. In Higher Education, as well as elsewhere, there is an increasing focus upon placing the learner at the centre of the process: student-centred learning. Indeed, recognising and addressing students' particular learning needs in an increasingly quality-driven higher education system has assumed even more importance in this world of change. Responding to individual needs and supporting learning are central concerns within this thesis.

Two themes that are especially concerned with student-centred learning are those of transparency of assessment and learners' involvement in the assessment process. Encouraging students to become involved in assessment mechanisms is not particularly new: Brown & Dove (1991), for example, report a number of initiatives involving peer assessment in UK Higher Education institutions that were taking place before 1991. Previously an exclusive club of which only academic staff could be members, assessment is now transforming itself into a transparent system that encourages others to join. Students, whether being invited to judge the work of their peers on the basis of assessment criteria, or in contributing informally to the development of module assessment, or through their institution's course evaluation and quality procedures, are encouraged to take advantage of the opportunities, benefits and also responsibilities for their involvement in these areas. Moreover, the considerable benefits that may be obtained by both students and staff through involving students in peer assessment in higher education, such as increased autonomy and responsibility, have already been extensively documented (e.g. Falchikov, 1998; Dochy *et al.*, 1999).

Of course, such benefits will not always occur when introducing a programme of peer assessment. A reminder of the pitfalls that may be encountered when using student peer assessment appear, for example, in the cautionary commentaries provided by Brown (1998) and Ballantyne *et al.,* (2002). These articles underline the potential dangers of using an assessment technique where the operational experience of tutors is inadequate or the rationale for its use is unclear to its participants.

My interest in the possibilities of using peer assessment to support student learning arose from an involvement in the teaching of undergraduate popular music courses during 1999. Noticing an apparent lack of peer assessment studies in the HE curriculum area of popular music group work, I became attracted towards exploring the use of this technique in students' band rehearsals and in the learning opportunities that take place in such contexts. During this time, I was involved in teaching two courses offered at Barnsley College (validated by the University of Sheffield): *BA (Hons) Popular Music Studies*, and *BA Extended Degree in Popular Music Studies* (a one-year HE foundation programme). Each course had a practical and real-world vocational ethos with many of the course learning outcomes involving practical group music-making experiences. Indeed, I considered that the popular music industry was characterised by group activities; working in groups, for example, typifies popular music performance, production, sound recording, distribution and promotion. For students of popular music (and similarly for the professional pop musician in the industry), group working is intrinsic to the discipline. Group work helps to prepare students for the real-world of popular music and, for this subject area, it can be considered as a natural setting for supporting learning.

A central activity of these courses involved group performance: rehearsing as bands followed by a performance on a stage. I was the students' tutor for these courses in which assessments were conducted at various times during the academic year. This allowed opportunities for me to develop peer assessment techniques that I considered were appropriate for assessing a band rehearsal, typically comprising four or five musicians, as well as their on-stage public performances.

This research developed shortly after my decision to explore peer assessment as a technique with which to assess performances given by groups of popular music undergraduate students. Peer assessment of group performances was introduced in the 1998/99 academic year and my ideas for this research were formed in September 1999.

The context for this enquiry, group-based rehearsals and performances of popular music linked with peer assessment, have together received little attention in the relevant literature and this is the educational setting of the research. The thesis is an attempt to expand our knowledge of these contexts by conducting an extended investigation into how peer assessment can support student learning. Many studies across a variety of curricula report various improvements across a range of activities as a consequence of participation in peer assessment. There is, however, much less known about how, why and which peer assessment activities involving band rehearsing and performing of popular music students can be effective in supporting their learning.

I became particularly interested in exploring whether the kinds of personal attributes that are often displayed during band rehearsals could contribute towards the development of a particular peer assessment system. The focus of this enquiry involves the contribution of peer-assessed personal attributes to student learning. It investigates what students might learn about themselves and others through peer assessment of their personal attributes during band rehearsing and performing. Personal attributes are an area regarded by Pearce (2000), for example, as being 'central to the education process' in developing learning. Also relevant is the issue he raises concerning '...[whether]... the musical ensemble awards marks to each other based on ...personal attributes' (p.46).

An action research approach was adopted in order to develop and improve my understanding and practice in this mode of assessment. Arising from this was the identification of a number of key issues in which particular variables involved in peer assessment were examined. These included: deciding on the type, purpose and sequence of peer assessment activities; the choices and decisions that may be encountered within a peer assessment system; investigating how peer assessment activities might impact on honesty, risk-taking, trust, confidence; the significance of all of these for student learning. Studying other variables such as gender, age, type of instrumentalist and prior experience, although interesting, was not considered to be feasible within the scope of this enquiry. Further details about the parameters of this research can be found in Chapter Two and Chapter Three.

This research is distinctive as a consequence of its chronological approach in which peer assessment is examined over a large number of rehearsing and performing cycles. If so, it contributes to an area of peer assessment literature that appears to have been neglected: the use of peer assessment in a group work learning context over a period of time (i.e. five years).

To summarise, the principal subjects involved in this research were undergraduate students of Popular Music involved in band rehearsing and performance. The majority entered the course from a Further Education background having obtained a National Diploma in Popular Music or Music Technology. A number of others possessed A Level qualifications and the remainder, comprising around one quarter of the total, were mature students. Approximately three quarters of all students were male and nearly all on the course had substantial prior experience as performers in bands. Further details about students' profiles can be found in Section 3.1.

An outcome of the research is the development of a new process model of peer assessment for group-work in music that promotes student learning through peer assessment in rehearsing and performing. The model is centred on students' active engagement in peer assessment, which involves students in developing assessment criteria and feedback that supports learning, within and informed by social constructivist pedagogy. Integral to the model is the development of a distinctive approach to peer assessment in which peers determine personal attributes for individual band members; this is a key construct that was explored during the enquiry and is discussed in the later chapters of the thesis.

Although the enquiry is situated within the context of popular music group rehearsing and performing, it is nevertheless expected that its findings and recommendations will have implications elsewhere in, for example, comparable performance-based or creative curricula.

The outline of the thesis is as follows:

Chapter One is a survey of the literature relevant to the thesis. This includes an examination of studies involving peer assessment, group work and music.

Chapters Two and Three develop the methodological underpinning of the thesis. The second chapter introduces the various methodological questions, considerations and assumptions that are implicit in this enquiry. The third chapter explains the research methodology.

Chapter Four considers the development of rehearsing and performing cycles that are an integral feature of the action research employed in the investigation.

Chapter Five presents the analysis and interpretation of the qualitative and quantitative data arising from the study.

Chapter Six develops a discussion about these findings and what they mean for student learning and peer assessment of student bands in group rehearsing and performing. A new model of peer assessment for this context is proposed.

Chapter Seven concludes the study. It discusses the implications of the research in other contexts and what contribution it makes to our knowledge and understanding of learning, teaching and peer assessment.

Chapter One: Literature Review

1.1 Introduction

1.2 Peer assessment and learning

1.3 Group work and peer assessment

1.4 The process of peer assessment

1.5 Peer assessment literature in the HE music curriculum

1.6 Epilogue: the claims and the pains of peer assessment

1 Literature Review

This chapter begins with an introduction that provides an overview of the key themes to be encountered in the literature on peer assessment in HE, followed by an exploration of those themes that have particular relevance for this thesis. Peer assessment can be a difficult and contentious topic in education, as indicated by the complexity of issues that are apparent when surveying the published literature. The purpose of providing an overview, at the outset of this chapter, is to indicate the range of issues and themes that are involved, in order to allow a contextualisation of these in the curriculum area of popular music.

1.1 Introduction

The use of peer assessment as a technique with which to assess the work of HE students is not regarded these days as an area of education that is especially new or novel. Indeed, much valuable work in exploring the use of this assessment technique was conducted prior to 1990. A number of important contributions have appeared in the literature since that time, however, that add significantly towards the development and application of sophisticated peer assessment schemes in Higher Education. A crucial element of this work was, and still is, the establishment of a rationale and its broad principles as a prerequisite for the introduction of peer assessment.

Developing a rationale for peer assessment and making sense of its basic principles are, of course, fundamental to any consideration of whether to employ this mode of assessment in a course programme. Much of the literature that has sought to develop a rationale for peer assessment has usually examined this in terms of the kinds of knowledge, understanding, abilities, skills and so forth, which were thought implicit in the activity; surprisingly few papers examine peer assessment in terms of its appropriateness to particular learning activities. Given the importance of understanding the nature of learning activities, and their potential for peer assessment, it is perhaps surprising that this relationship does not always attract explicit comments in the literature.

General principles of peer assessment together with the practicalities of introducing and managing its process mechanics have been the subject of many studies. Examples of a

number of these are described in the various sections of this chapter. These studies occasionally offer suggestions for tutors who may be thinking about introducing peer assessment into their own practice, along with practical tips drawn from particular experiences, such as how students might receive preparation or training in the process.

In addition to the literature that includes a focus on general principles or rationales for peer assessment, there are also a large number of studies involving aspects of peer marking, such as examining the accuracy of students' marks, compared with those of their tutors. Another key theme to be found in the literature concerns methods by which peer assessment criteria are established. The chosen method of devising criteria for peer assessment is often considered as an important part of the peer assessment process. The extent of the care that was taken in developing criteria for students to assess their peers is revealed in a number of studies.

Several approaches to establishing peer assessment criteria described in the literature involve criteria jointly set by students and tutors. Involving students in the setting of peer assessment criteria is often explained and justified in terms of providing opportunities for students' learning experiences. Examples of studies reporting this are listed in Table 1.1. (p.64). Also, it is argued, involving students in this part of the process improves the transparency of the peer assessment system itself. Perhaps less attention, however, is given to difficulties that might arise from empowering students in this way: for instance, poorly expressed or limited criteria that do not test the appropriate learning outcomes of the assignment or that fail to address the required academic standards. Examples of studies reporting these difficulties are listed in Table 1.2. (p.66).

There are also a number of studies that integrate peer assessment with other forms of assessment, particularly that of self- assessment. Indeed, a survey on assessment literature in these areas suggests that self and peer assessment, although separate and distinctive assessment methods, enjoy a relationship of an overlapping and mutually reciprocal character.

There appears to an increasing emphasis upon more focused aspects of peer assessment practice. There are studies, for example, that examine the processes, practicalities and

pedagogical contexts of group peer assessment; strategies for the stimulation of learning; and case studies of practice arising from particular areas of the curriculum that might have applications elsewhere. Examples of these can be found in Bryan (2004), as discussed later in Section 1.3.5.

These, then, are some of the key themes that may be encountered when surveying the literature of peer assessment in HE. At this point, it may be timely to clarify what is meant by group peer assessment and to explain its defining characteristics. Topping (1998) has provided a definition of group peer assessment that has also been adopted by others (for example, Bilgin & Fraser, 2007):

> an arrangement in which individuals consider the amount, level, value, worth, quality, or successfulness of learning outcomes or products contributed by each individual member of the group. (Topping, 1998, p.250).

His definition suggests a link between the activity of peer assessment and opportunities for learning, a theme that occupies a central position in this thesis. Indeed, the term 'peer learning' is preferred by some in order to indicate 'learning' rather than 'marks' as the key priority. The use of the term 'peer assessment' in my study, similarly, should be understood as having a wider application and meaning. The term 'peer assessment' in my study follows the account of its meaning and application that is offered by Brown (2000): 'Generic term for assessment of peers by peers. Also used to describe peer support, peer feedback, peer learning, peer feedback marking.' (p.2)

Before continuing with an exploration of the literature it may be helpful to be reminded of the research focus as stated in the Introduction: the contribution of peer-assessed personal attributes to student learning within the context of band rehearsing and performing activities. This chapter is organised in the following sections.

1.2 Peer assessment and learning
1.3 Group work and peer assessment
1.4 The process of peer assessment
1.5 Peer assessment literature in the HE music curriculum
1.6 Epilogue: the claims and the pains of peer assessment

1.2 Peer Assessment and Learning

The flavour of this section and its significance, not only in this literature chapter but also in this thesis, can be summarised in the description by Gibbs (2006) about the point and purpose of peer assessment. His view also articulates a belief shared by many about what should be at the centre of peer assessment.

> Much research on self- and peer assessment appears to be obsessed with the reliability of student marking in the hope that student-generated grades can substitute for teachers' grades and save the teacher a great deal of work. If you go to enough trouble, students are indeed capable of reliable marking...but this completely misses the point. What is required is not more grades but more learning. (p.27)

Expressing a similar sentiment a decade earlier, Parsons & Drew (1996) welcome a change of emphasis in assessment from its use as a 'control mechanism' to 'assessment to aid learning' (p.72). Research involving peer assessment has begun to address the learning benefits that may be obtained through the employment of this method of assessment. This section identifies and focuses upon key developments in peer assessment that the literature suggests offer particular scope and opportunities for supporting learning.

There is increasing focus today on student-centred learning experiences and the importance of providing opportunities for students to be actively engaged in their education. The purpose of this section, therefore, is to survey and investigate the support for learning that has been claimed in the literature.

David Boud and Nancy Falchikov have each written extensively about assessment, including peer- and self-assessment. In their joint discussion of the need to align assessment with long-term learning (Boud & Falchikov, 2006), many important points are raised. The two main purposes of assessment, they assert, are firstly to provide certification of achievement and secondly 'to facilitate learning' (p.401). As a consequence they argue that equipping students for their journey of 'lifelong learning necessarily involves preparing them for the tasks of making complex judgements about their own work and that of others' (p.402). If so, then providing opportunities whereby students may acquire and practise their judgemental and evaluative skills is critical towards the development of a rationale for peer assessment. Graduates are likely to be

making judgements about what is good or not in their workplace and about their colleagues' teamwork skills and personal attributes. They are less likely to be preparing for examinations or producing academic written work.

1.2.1 Self assessment and peer assessment

Self- and peer-assessment is often coupled in the literature and the inclusion of a self-assessment element appears in a considerable number of peer assessment studies. In order for peers to be assessed by peers, prior skills in self-assessment are thought, by some, to be a pre-requisite (Pope, 2005, p.52). For peer assessment to be effectively utilised therefore, Pope believes that students should have undergone some preparatory self-assessment or at least engage in an appropriate self-assessment activity within a peer assessment assignment. If students can be engaged in self- and subsequently peer-assessment, as Boud (1991) and Ellis (2001) observe, this can allow a mutual and reciprocal exchange of information

However, if self-assessment is a pre-requisite to students obtaining a meaningful peer assessment experience, it raises a question about why much of the literature on peer assessment does not report prior self-assessment experiences. It is an issue, also, when involving first year undergraduates (advocated by many as being the most appropriate place for introducing peer assessment) where self-assessment experiences may be somewhat limited. If a tutor has no knowledge about students' involvement in self-assessment activities elsewhere on the course, it may, as a demonstration of good practice, be prudent to include a self-assessment element within the peer assessment experience as advocated by Boud. Students' self-assessment activities in this research are explained in Chapter Three and Chapter Five.

1.2.2 Some perceived benefits of peer assessment

A survey of the literature of peer assessment is not short on claims made for its various benefits as an assessment technique, particularly in areas involving learning. It is the belief of Prins *et al.* (2005) that courses which include peer assessment can have multiple beneficial effects in terms of learning goals. They label *content related* improvement as the first learning goal and the *acquisition of peer assessment skills* as a

higher order learning goal. For example, research by Sluijmans *et al.* (2002) concluded that students significantly improved their end products (the creation of study tasks) under peer assessment conditions and this was interpreted as 'a positive effect.... in the content domain.'(p.451). Prins goes on to suggest that examples of the higher order peer assessment skills might include deeper reflection by students on their own behaviour and performance, increased confidence in their abilities to perform to specified criteria, and awareness of the quality of their own work. Indeed peer assessment can help students learn how to make judgements about performance that are important for them in becoming 'reflective practitioners' (Schön, 1987, p.175).

Claims that are of a general nature such as these appear repeatedly in the literature. For example, Ballantyne *et al.* (2002) indicate a number of areas in which the literature suggests students' learning can improve; briefly, these include encouraging students to consider the objectives and purpose of the task, deciding upon the quality of work, better understanding of assessment processes, and promoting reflection upon assessment tasks and improving self-confidence. Falchikov (1998), in a survey of various peer assessment initiatives, presents an impressive list of perceived benefits to students. These include: group working and interpersonal skills, transferable skills, organisational skills, listening skills, speedy and extensive feedback to large numbers of students, assisting students with moderating their learning behaviour and improving performance, enthusiasm and motivation, ownership, increased confidence, improving understanding, reflection and intellectual development. Her claims are expanded later in the chapter and several of these are incorporated in Table 1.1. (p.64)

Gibbs (1999) believes that an important consequence of peer assessment resides in increasing students' responsibility to themselves and to others due to the social pressure that is brought to bear. This is sometimes described as developing 'self-regulation' (Boud, 1986; Boud, 2000; Nichol & Milligan, 2006).

There are a number of reported learning benefits that are related specifically to students' improved reflection upon their own work and abilities including Somerville (1993), Anderson & Freiberg (1995), McNamara & Dean (1995), Longhurst & Norton (1997), Orsmond *et al.* (2002), and Keppell *et al.* (2006). Reflecting upon feedback as a result

of peer assessment is thought to be a particularly important part of the learning process. The literature involving feedback is specifically considered in Section 1.4.

A large number of studies exist in the major reported literature of peer assessment in HE (over a period of around two decades) that report a wide range of learning benefits in their findings. This accumulation of experience and knowledge provides some compelling arguments for the inclusion of peer assessment activities within higher education. However, as a mode of assessment, while appearing to offer much in providing opportunities for supporting learning, it is not without its problems. Many practitioners draw attention to the risks involved and warn tutors to proceed with caution when adopting this type of assessment. Indeed the views of Purchase (2000) and Kwan & Leung (1996) that peer-learning opportunities outweigh risks of inappropriate assessment, especially if only a small proportion of students are involved, may appear somewhat contentious.

1.2.3 What are the areas that are being assessed?

Falchikov (1998, p.13) explains that the objectives and purposes of peer assessment can be divided into three basic areas: assessment of the *product* (for example written, oral, and visual in which posters seem to be popular); assessment of the *process* (for example the activities of the group, including students' 'explicit reflection' (Cowan, 1998); assessment of professional skills. This last area, she observes, is not frequently done. Similarly, Pope (2005) identifies two divisions: *performance/process* (for example discussion or group project) and *product* (for example computer projects or posters). He points out that while performance/process studies provide the benefits of involvement and understanding, they cannot be used usefully for the summative grading of students' discipline knowledge. It is rather, claims Pope, the product type of assessment that is more suited for this purpose.

Although not explicitly articulated in Pope's paper, a consideration of such divisions serves to remind tutors of the need to be clear about the purpose of employing peer assessment. This includes setting appropriate intended learning outcomes, and identifying which type of activity – process or product (or possibly both) – is considered best suited to facilitating that kind of learning. For Race (1998, p.114), the two major

aims of peer assessment are to allow students to learn from each other's successes and from each other's weaknesses. Applying assessment to peer learning which involves students' strengths and weaknesses in this way raises some interesting questions. Assessment involving the somewhat personal area of students' strengths and weaknesses requires great care and sensitivity. An important point here, arising from Boud & Walker (1993), is the distinction between the person, who is always valued, and particular acts or specific activities, which may be subject to critical comment. These areas, involving strengths, weaknesses, skills, qualities, and learning contexts, are of considerable importance for this thesis and underpin Chapters Six and Seven.

There is the issue also of whether learning benefits, arising from students' successes or weaknesses, develop in equal measure across both *process* and *product* types of assessment. For example, would directing peer assessment towards the area of students' strengths and weakness be suggestive of a focus upon process aspects or be for formative purposes? A focus on process is also important for Biggs (2003) and for his 'constructive alignment' framework, which is discussed in the next section. Similarly, would it be possible to separate the product that has been created from the activity of creation when assessing strengths and weaknesses? These are among the questions that are investigated in Chapters Two, Six and Seven in the thesis.

Students possess skills, qualities, capabilities, attributes of varying strengths and weaknesses. The context in which these are displayed is crucial. The theme of this section is learning, and what is being assessed. Fundamental to assessing students' strengths and weaknesses during an activity would seem to be the provision of a suitable and meaningful context in which learning can occur, whereby students are able to develop their skills, qualities and attributes. Discussing appropriate contexts that may be amenable to the peer assessment and the kind of learning being assessed recalls the *Situated Learning* theories of Lave & Wenger (1991). Two principles of situated learning are that knowledge needs to be presented in an authentic context (settings and applications that would normally involve that knowledge) and that learning requires social interaction and collaboration. Their work is also discussed later in this chapter.

Skill-based learning is emphasized in a peer assessment model developed by Sluijmans & van Merrienboer (2000). Their model is divided into three skill-related activities commencing with defining the assessment criteria appropriately for product or process.

The next activity consists in making judgements about the peer performance, reflecting and identifying strengths and weaknesses and writing an assessment report. The third and final activity is that of providing feedback for future learning. They propose that training programmes should be centred on developing these skills. Their model, however, seems to lack the provision for developing student experiences or training in self-assessment skills, as advocated by Pope (2005). Self-assessment skills are integrated in the process model that is proposed in Chapter Six.

There are a small but increasing number of peer assessment studies (Knight & Yorke, 2003; Bloxham & West, 2004; Sambell *et al.*, 2006) that acknowledge the influence of the teaching and learning ideas of John Biggs (2003) in his book *Teaching for Quality Learning at University*. It is appropriate at this point to outline something of the significance of Biggs' work, albeit briefly.

1.2.4 The 'constructive alignment' framework of Biggs

Biggs uses the term 'constructive' to explain that learning is constructed by students as they interact with their experiences in the world. Tutors can support student learning through aligning teaching method and assessment to the learning activities (p.11). Engagement with learning activities can lead to deep learning and involves a focus on process features rather than content (p.16). The development of graduate and professional skills depends upon the acquisition of three kinds of knowledge: relevant declarative or relevant subject; procedural or knowing the necessary skills; conditional knowledge or knowing how to apply acquired knowledge. Constructive alignment (p.25) supports learners in developing and acquiring all three to produce fully functioning knowledge. Biggs describes how assessment should be appropriate for and aligned with students' active learning experiences. As such, he develops the concept of 'performance assessment' (p.156) to enable an active demonstration of knowledge. Included in appropriate 'performance assessment' activities are problem-solving and group projects.

It is possible to recognise something of the flavour of peer assessment in constructive alignment. For example, each of the following constitute activities might exemplify constructive alignment principles to some extent.

Firstly, Biggs' emphasis on active learning experiences might be exemplified by engagement in groupwork interaction, an activity that is typical of many peer-assessed assignments. Secondly, involvement in the peer assessment mechanism itself could be described as an active learning experience as it normally requires a number of participants. Thirdly, the creative or process element of a peer-assessed activity is often identified as a focus in peer assessment (typically when a 'contribution' type criterion is employed). Finally, tutors often devise peer-assessed assignments with real-world activities in mind. Constructive alignment can support preparing students for their existence in life beyond university, through the development of deep approaches to learning. If so, then in these areas at least, do 'constructive alignment' and groupwork peer assessment, founded on constructivist pedagogy, share common principles?

Biggs adopts a pragmatic approach and relates theory to practice in a practical manner. It is not difficult to relate his ideas of engagement and alignment to rehearsing and performing contexts. This is revisited in Chapter Two in which the popular music course in this research exemplifies aspects of 'constructive alignment'. An 'active engagement' paradigm is also proposed in the final chapter.

1.2.5 The social constructivist assessment process model of Rust, Price & O'Donovan

Other assessment frameworks are emerging that have relevance for peer assessment. One of these is the social-constructive assessment process model (Rust *et al.*, 2005; Price & O'Donovan, 2006; Rust, 2007). Acknowledging the influence of Gibbs (1992), Brown *et al.* (1994), Biggs (2003) and others, the authors develop a diagrammatic model consisting of two 'parallel ongoing cycles, one for staff and the other for students' (Rust *et al.*, 2005, p.236). Rust, Price and O'Donovan emphasize that the two cycles are not separate but rather are two halves of a dynamic system with common understanding being shaped within a 'community of practice'. One half is concerned with the discussion and development of explicit criteria, while the other emphasises students' active engagement with criteria, feedback and self-assessment. Appropriate diagrammatic connections are made between each half. The authors advocate that the two cycles should occur at least twice on the module or course. The second visit would allow consideration of the work and feedback as part of the tutors' initial engagement

with the criteria in the next cycle. The process is cyclical therefore, with each successive cycle informing the previous.

Although the model addresses assessment in general, it may also have relevance for peer assessment, particularly where peer assessors are actively engaged in the development of criteria and equally so in the feedback process. This model revisits a number of key themes that have been explored in this section of the chapter. These include: involving students in the setting of assessment criteria; the importance of feedback and self-assessment; engagement in active learning contexts illustrated by critical models of teaching and assessment developed by Biggs and Rust *et al.* It is an opportune place at which to conclude this peer assessment and learning section of the literature.

1.3 Group Work and Peer Assessment

The literature on group-based work in Higher Education is extensive, as student group work activities can be found in every subject discipline. The purpose of this section is to explore an appropriate and selective literature of group work that is thought to have a particular relevance for either peer assessment or popular music group activities. Establishing a focus for this survey into group work and peer assessment is not a simple task; such is the wealth of available material and the challenge involved in evaluating the relevance and potential of each item, in terms of a popular music context. Nevertheless, it is hoped that the literature search conducted in this area produces a satisfactory breadth, depth and exploration of the relevant issues. This survey of the literature of groupwork and peer assessment is organised into the following areas considered relevant to the thesis: *Theoretical models of group work*; *Cooperative and collaborative learning in group work; Determining group membership;* The HEFCE/FDTL *Assessing Group Practice project; Characters in groups: shooting stars and free riders.*

1.3.1 Theoretical models of group work

Working effectively as a member of a group is integral to many curricular areas of HE as well as a part of lifelong learning. This is particularly so for popular music performance, where in both real-world professional and undergraduate contexts, group

work is a typical and distinctive activity; for example, pop musicians working together as a team, combining their creative and musical abilities in order to accomplish effective rehearsals, performances or recordings.

Group difficulties caused, for example, by deteriorating interpersonal relationships, creative tensions, accusations of musical incompetence, or lack of commitment, are typical problems that are experienced at one time or another during band rehearsals. Undergraduates who rehearse in a band as a part of their course usually experience tension within their group at key moments for a variety of reasons. Within the literature on group work, there exist a number of theoretical models that attempt to explain the developing tension and dynamics within groups over a period of time. This literature may have some relevance for studies of popular music bands and their rehearsing.

For example, popular music students would certainly recognise from experience, particular characteristics of their rehearsing in relation to the theoretical phases of group development described by Tuckman (1965) and Tuckman & Jensen (1977). Their work suggests that all groups pass through four sequential stages of development. These stages may be longer or shorter for each group, but all groups will need to experience them. They are *forming*, *storming*, *norming* and *performing*. Peer assessment may have some impact upon bands as they pass through the stages that are characterised by Tuckman & Jensen.

Certainly for popular music students, where band tensions and disagreements between individual members do appear, these often occur during the intermediate part of the rehearsal period: the *storming* stage. During the outset of a group performance assignment, bands typically exhibit levels of motivation through being absorbed in planning the rehearsals: group membership, instrumental roles, deciding on the set, and so forth. Similarly, during the closing stages of band rehearsals, the lever of a performance in public often has an effect of increased motivation to perform well. For some bands, however, the *storming* stage becomes protracted, leaving little time for them to reconcile their conflicts necessary for proceeding to *norming*. If so, would the use of peer assessment in the rehearsal process assist bands to proceed more speedily through the difficult *storming* stage, or might it exacerbate tensions?

The self-perception inventory of Belbin (1981) was developed as a means of giving group members a way of assessing their best team roles that, arguably, resonates with contemporary popular music group performance. Indeed, Bryan (2001) suggests that this inventory could be adapted towards developing effective group behaviour in performance activities through student peer-observation. In a more recent article, Bryan, in Bryan and Clegg (2006), explores the application of 'task' and 'group' maintenance ideas (Bales, 1970; Douglas, 1976; Jacques, 2000) in the context of problem based group work. She provides examples of possible group maintenance criteria including listening to others, enabling shy members to contribute, and techniques for dealing with unproductive disharmony within the group.

Hartley (1997) presents a theory of group communication which might have some relevance to group work, entitled 'An Integrative Model'. Hartley brings together three levels of analysis needed to comprehend group interaction: social and cultural background, tasks and procedures, interpersonal underworld. Firstly, the activities implicit in band rehearsal and performance of popular music – preparing for a gig in front of an audience at a public venue – is one that is most definitely situated within social and cultural contexts. Secondly, Hartley defines the 'surface behaviour' level of groups in terms of 'task' and 'procedures' (not unlike, perhaps, the task and group maintenance functions described in the previous paragraph). Finally, and perhaps the most interesting of Hartley's levels, is that of the 'interpersonal underworld'. This refers to the 'pattern of likes, dislikes, admirations, resentments and all other emotional attachments which exist between the group members' (p.29). It is not difficult to equate this description with the popular portrayal of rock bands and tensions between individual musicians often portrayed in the media, using such metaphors as 'differing artistic and creative directions', 'irreconcilable professional differences', 'a clash of egos' and so forth. The development of a peer assessment process model in chapter seven arises, in part, from issues related to the 'interpersonal underworld' of undergraduate bands.

Although several other influential studies of group work exist (for example, Jaques, 2000) it is the intention of this section of the chapter to suggest examples where an interface may possibly exist between group work theory and popular music bands.

Group work and team projects can be problematic in a Higher Education setting. Heywood (2000) believes that one reason why several curricula shy away from group based tasks is that many tutors believe that it is impossible to assess them: yet the ability to work effectively in a team is a quality that is highly sought by many employers. The next section examines the literature that addresses such prized qualities that may be developed in cooperative and collaborative group work.

1.3.2 Cooperative and collaborative learning in group work

Group work allows opportunities for collaboration and cooperative learning to develop. Cooperative and collaborative learning in groups is becoming an increasingly common approach in many curricula activities. The terms cooperative learning, collaborative learning, peer learning and group learning are often interchangeably used to define a process by which students work jointly in small groups to accomplish an educational task (Boud *et al.,* 1999; Boehm & Gallavan, 2000; Gupta, 2004). The virtues and benefits of working cooperatively in groups are extolled by many and the perspective of the student who is an 'academic loner' in classrooms is very different from those working in cooperative learning academic teams (Stahl and VanSickle, 1992). As well as guarding against the danger of isolation, particularly in a student's first year, Hunter (2006) suggests that collaborative activities can also serve as a valuable retention strategy. The Higher Education Academy subject centre for Dance, Drama and Music (PALATINE) contains a number of recent collaborative initiatives and associated research. Kleiman (2004) presents a number of benefits from small-group learning in a collaborative environment including celebration of diversity; acknowledgement of individual differences; interpersonal development; actively involving students in learning; more opportunities for personal feedback. The relevance of this approach to learning and skill development for the activity of band rehearsing is sufficiently obvious as to hardly be in need of further comment here. However, 'celebration of diversity' in the context of a group of students brought together to form a band could equally be interpreted as celebrating different musical tastes across popular music genres.

Stahl (1994) in an informative overview of cooperative learning, suggests a number of essential requirements to be met in order for groups to succeed within a cooperative learning framework. One of these requirements involves providing a setting for enabling

positive social interaction, in which students can develop interpersonal qualities such as leadership, trust building, conflict management, constructive criticism and compromise. These qualities, apart from sharing a certain likeness to Belbin's theories of team roles (Belbin, 1981), are also relevant to band rehearsal dynamics, particularly in the area of resolving difficulties.

Cooperative learning groups often employ self or peer-assessment. In a physical sciences course involving cooperative learning which employed peer assessment, Gupta (2004) concluded that it encouraged students to make the most of their opportunities The study by Prins *et al.* (2005) of a computer supported collaborative environment employing formative peer assessment, emphasises the need for fruitful interaction, that students must be individually accountable and that positive interdependence between group members has to be stimulated. Establishing face-to-face environments was particularly difficult to achieve in that specific curriculum activity.

Face-to-face interaction was also another requirement of Stahl (1994) in which he advocates that students need to position themselves where they are able to face each other for direct eye-to-eye contact. Popular music students in their band rehearsals adopt such positioning naturally. Facing each other in a rehearsing room, often with the drummer as a focus, facilitates musical interaction between each musician: vocalists, instrumentalists and drummer. It especially allows non-verbal communication during the music making (such as the nod of a head to indicate the start of a guitar solo, or the drummer emphasising a song's structure using exaggerated movements). In fact, so accustomed are students to rehearsing through cooperative positioning that it can become problematic when transferring on to a stage. Student bands who have not rehearsed as they would be in front of an audience - the vocalist down stage centre, instrumentalists in a linear placement from centre left to centre right, with their drummer positioned upstage and behind them, rapidly discover that cooperative eye-to-eye contact becomes difficult. Face-to-face interaction involving giving and receiving peer feedback, an area that is of importance to this thesis, is discussed, in Chapter Five.

Face-to-face interaction and collaborative skills are often a part of every day life in the real world. Race (1999) reminds us that many employers value cooperative and collaborative skills highly and there should be a balance between collaborative and

individual work in higher education. He also points out certain dangers when trying to measure collaborative work objectively and, somewhat teasingly, suggests that 'if we can assess it, it probably isn't it!'(p.68). In continuing, he suggests a viva examination as a means for assessing students' collaborative activities.

Cooperative and collaborative work is not necessarily identical to group work. Just because students are placed in a group it does not necessarily follow that interaction and cooperative learning is always taking place (Johnson *et al.*, 1993). The term 'jigsaw activities' Stahl (1994, p.24) is used to indicate separate tasks undertaken by individuals, similar to the pieces of a jigsaw that become necessary to complete the picture. Such activities cannot claim to be cooperative learning tasks. In the same way, student bands may decide to assign particular tasks to individual band members: obtaining recordings of chosen songs; acquiring guitar tablature music of the songs; practising the individual instrumental parts; and organising the schedule of rehearsals and booking rehearsal rooms and so on. Although these activities, valuable as they are to the group as a whole, might constitute cooperative group tasks, none would be described as examples of cooperative learning experiences. Hunter (2006) similarly describes group seminar presentations in music, consisting of individuals who simply take it in turns to read a part of the seminar paper, as not exemplifying true collaborative work.

The two case studies presented by Sambell *et al.* (2006) (involving first year undergraduates on an interdisciplinary humanities degree and a second year course in engineering) state that an emphasis was placed on collaborative activities in order to 'draw on the social learning potential of the student group' (p.159). Although there were some very interesting and useful findings regarding how these students developed procedural and critical autonomy, it is unclear how the tasks, as described, constitute genuine examples of *collaborative* learning.

In the context of the performing arts, cooperative and collaborative learning activities, although often present, are rarely a focus due to the emphasis on individual objectives and performance (Bryan, 2006). She suggests that, within group work, students might develop and acquire important collaborative skills such as communication, negotiation, self-initiative, resourcefulness and conflict management – all highly relevant and useful

qualities for working in bands. Hunter (2006) suggests that musicians are probably at an advantage over other students from certain other disciplines in that from an early age they have usually been involved in collaborative music making activities, such as choirs and classroom instrument group work.

This section has, at this point, identified a number of points arising from group work literature that have relevance for peer assessment of group work musical activities. One of these was the 'interpersonal underworld' within groups and their attendant emotional tensions; another concerned the opportunities in group work for cooperative learning and the requirements of trust, leadership, conflict management, constructive criticism and compromise; a third involved face-to-face interaction and the development of collaborative interpersonal skills.

The success of a group can depend on such factors as its membership, its size, the characters within it and the role of the tutor in facilitating its functioning. The next section explores some of the literature relating to these and to similar considerations.

1.3.3 Determining group membership: its basis and group size

An important decision required early in the process is one that concerns the student groupings themselves: determining, for example, the appropriate group size and deciding the basis for selecting individuals to be members of a group. This section confines itself within the literature on peer-assessed group-based studies.

Given that deciding the membership of a group is an important factor for students, tutors and the nature of the task, it is surprising the number of studies that do not explain the basis for how such decisions are reached.

Biggs (2003) draws attention to the important pre-requisite that students should have sufficient background knowledge or skills in order to be able to contribute. There is, he suggests, no optimal group size, as this is dependent upon factors such as the nature of the curriculum activity, its intended learning outcomes and, as Bryan (2006) also believes, the dynamics of the class.

Exley & Dennick (2004) propose that the 'optimum size' for small group teaching 'is between 5 and 8 per group' (p.2). Five-piece bands were, in fact, the most frequently occurring group size in this enquiry. Three-piece bands, however, are also standard group sizes appearing in popular music; countless professional acts comprise a singer-guitarist, bass player and drummer. Examples of highly successful three-piece rock bands, encompassing a period of over forty years, include *The Police*, *Cream* and the punk outfit *Alkaline Trio*.

Undergraduate group work involving popular music performance is, perhaps, no different from many other areas of interpersonal cooperation in which individuals may exhibit differentiated levels of personal contribution. One distinctively musical difference, however, resides in the formation of groups for the purpose of performing music. This context implies the availability of appropriate instrumental sounds and timbres through which the creation of popular music can become meaningful and communicable. Bands typically contain drums, guitars, bass, keyboard and vocals, although, of course, other instruments may be also required. However those five instruments (including the voice as an instrument) are typical of a basic band line-up. Group membership for this research was decided on this basis and is explained further in chapters three and four. Instrumental versatility, or being proficient on more than one instrument, is common among many pop musicians and often allows a band the option to decide on various instrumental permutations for the band, with its rhythm section (drums, bass, rhythm guitar) usually at its core.

This musical consideration has implications for determining the membership of bands, for one based on a randomised selection of instrumentalists would be inappropriate. A more usual method would be one in which bands are formed on the basis of including individuals with expertise on each of the instruments to be found in a basic band line-up. Usually three-, four- or five-piece bands are adequate for this purpose. For groups comprising more than five musicians, the more likely it may be that 'social loafing', a term used by Biggs (2003), will occur. The particularities of determining group membership in a popular music context perhaps shares something with other disciplines where teams depend, for their proper functioning, upon individuals who have specialist skills or abilities: for example in *sport studies* (bowler, batsmen, wicket-keeper), *film and media* (cameraman, sound engineer, lighting technician) and so forth. In courses

where students receive specialist tuition as an integral part of the curriculum (for example individual instrumental lessons) this often has a bearing on how groups are formed.

1.3.4 Who decides the membership?

It is interesting to consider the merits of whether tutors or students should decide the membership of groups. Information about how group membership is determined is often omitted in reports, which makes it difficult to classify peer assessment studies into, for example, tutor-determined groups, student-determined groups or randomised selections, this.

The popularity of tutor-determined groupings may partly be explained on matters of principle: many academics and students share a belief, for professional and ethical reasons, that tutors alone should determine group formations. Another factor in favour of tutor-determined groups is that selection can be informed by the tutor's knowledge of the class as a whole as well as individuals. If facilitating effective group dynamics is regarded as a key indicator in determining a group's potential achievement (Bryan, 2006), then the tutor armed with knowledge of the class is well placed to configure students groupings, as appropriate, in order to achieve optimum group conditions.

Although a number of peer-assessed studies report that students were allowed to decide the groups among themselves (for example, Lloyd-Jones & Allen, 1997), very few explain the rationale behind adopting this approach. Do students self-select upon musical ability, with the more able students attracting each other to form 'super-groups' in order to enhance the possibility of their achieving a high peer assessment grade? What were the implications for those groups who contain the remaining weak or poorly motivated students, and will these groups become dysfunctional? How do tutors accommodate unpopular students who, as a consequence of the self-selection process, are not chosen by any of bands? These are among some of the important questions in the literature that are often unanswered. This thesis revisits these and other similar issues in Chapter Three. Interestingly, the survey conducted by Lejk and Wyvvill (2001b) into group members' responses to peer assessment reported that the majority of students favoured choosing among themselves those with whom they would like to work.

Finally, some studies (for example, Lejk and Wyvvill, 2001a) adopt a randomised approach to the formation of groups. Although randomised groupings, for reasons explained earlier, were considered inappropriate for establishing bands, it may be argued, nevertheless, that randomised groups demonstrate transparency of process and tutor impartiality. If randomised selections suggest that the group task does not depend on specific skills or qualities of individuals, does it follow that there is no necessity for the tutor to possess knowledge of the individuals within the class? Randomised groups also ensure that students receive the experience of working with, and learning from peer colleagues whom they would not normally choose. Randomised grouping inhibits the occurrence of self-selected friendship groupings, and this might decrease opportunities in peer assessment for friendship marking.

In summary, there are three general approaches (i.e. tutor-determined, student-determine, randomised selection) to establishing group membership and each of these methods may have an influence on group dynamics and peer assessment. Given these implications, some studies in the literature have surprisingly little to say in either describing the chosen selection mechanism or the reasoning behind their choice.

Summary

This section on group work and peer assessment has surveyed a range of issues and questions arising from the literature. It introduces some important theoretical models of group work (i.e. Tuckman, Belbin, Douglas, Hartley) in order to explore their possible relevance and application to peer assessment, popular music and group rehearsing and performance. This was followed with an examination of a key argument in support of group work: providing opportunities for cooperative and collaborative learning experiences. The literature surveyed here has implications for the kind of cooperative and collaborative skills and qualities required for successful group work learning activities in music. Finally, a discussion developed from issues concerning the basis and principles of determining group membership.

At this point, it may be timely to examine an important source in the literature known as The Assessing Group Practice Project. This project led by Bryan (2004), involving several partner HE institutions, has resulted in a number of studies that are of relevance,

such as collaborative group work, group membership issues and peer assessment. The project can be regarded as a significant contribution to the literature of group work and peer assessment and deserves to be examined in some detail.

1.3.5 Assessing Group Practice Project 2004

This project was funded by HEFCE's Fund for the Development of Teaching and Learning (FDTL) in the period 2000 to 2004. It was undertaken by a consortium led by the Central School of Speech and Drama and included five other Higher Education institutions: University of London, Goldsmiths College; Dartington College of Arts; the Universities of Leeds, Middlesex and Ulster. The project aimed to ensure that assessment of collaborative activity within the performing arts could be demonstrated to be fair, robust and practicable across the sector. The identification and sharing of successful practice was seen as a key means to this end. The project culminated in a collection of papers published by the Staff and Educational Development Association entitled *Assessing Group Practice* edited by Cordelia Bryan (Bryan, 2004) and online at the Higher Education Academy subject network site for dance, drama and music: PALATINE.

Briefly, the project considered three overarching questions:
What are collaborative skills and how are they assessed?
Should we assess individuals or groups on projects involving collaborative work?
Can we reward creativity that goes beyond the stated learning outcomes?

The project offers some practical answers to these questions in the devising and presentation of a development workshop entitled *Peer Assessment in Group Work* suitable for tutors or students, along with case studies; 'how to' guides for students; reflections and commentaries on themes.

There are some interesting ideas and developments reported among some of the case studies. Of the case studies, there are four examples of group work and peer assessment that have particular relevance. The first example features a group-based theatre studies activity at the University of Ulster introduced in a year two course. The next, from the University of Leeds, describes the basis for group formation and the method employed

for feedback. A jazz performance activity from Middlesex University, although not involving peer assessment, illustrates the importance of a distinctive context through which students' personal qualities and attributes become significant. The final example, from the Central School of Speech and Drama, is interesting in how the assessment criteria are developed during the assignment.

BA Theatre Studies year two course at University of Ulster

Students selected their own groups in order to work through a mini performance of an example of political theatre practice. Each group drew up a set of criteria by which the piece might be judged. The pieces were presented as they developed and peer-assessed as a means of providing critical formative feedback. The tutor reported that the formation of groups had not proved problematic, but students were resistant to tutor-intervention in the selection and to any element of peer-assessment for grading purposes.

BA Theatre Dramaturgy year two course at University of Leeds

Peer tutoring and peer assessment were integrated into the teaching and assessment of the practical element of this module. The students prepared their presentations in groups but were assessed individually. The presentation element, although examined individually, involved the students working during the preparation stage in groups of three as performers or as technical assistants and operators. The students chose their own groups, decisions tending to be made on the basis of proximity of living accommodation, focus of study and friendship groups.

The students worked to tutor-given criteria. After a briefing, including discussion of the project and assessment guidance, the students were involved in a process of peer tutoring. Students were required to give feedback at specific points during the process to the other members of their group. It was delivered electronically and a copy sent to the supervising tutor. The tutor reported that students received this approach to the preparation and assessment of the practical submission very positively. It constantly drew students' attention to the assessment criteria and also made them aware of the support available to them from their own peer group. Students were relieved and surprised to discover the great lengths tutors go to in order to create a just and fair

system, which attempted to benefit the student by rewarding their learning and achievement.

The method of peer feedback requires some comment here. The use of email has the advantage of speed; speedy feedback is acknowledged as being good educational practice. However, why was there a requirement for these emails to be copied to the tutor? Such practices do not sit easily among the concerns (appearing elsewhere in this chapter) of McDowell & Samball (2005) and O'Donnell & Topping (1998), about authority figures accessing detailed feedback.

BA Music (Jazz) year two and three course at Middlesex University

The assessment process recognised the nature of jazz performance insomuch as it cannot all be prepared in advance. Students are placed in situations to which they must respond at that time and in collaborative circumstances. For their assessment, students take part in group performances that are partly improvised 'in the moment'. Because the music is composed as part of the performance, effective group collaboration is necessary to produce a successful performance product. Consequently there is a need to be adaptable and responsive to the ideas and contributions of others and to assume leadership and supportive performing roles. Those students who are able to 'pull others along' in the collaborative situation are assessed accordingly.

Jazz performance provides an interesting example of collaborative activities. Clearly there is a link between this activity and particular qualities and attributes of the students as a consequence of the distinctive learning context.

BA Theatre Practice year one course at The Central School of Speech and Drama

A group assessment system used at Level One of the BA (Hons) Theatre Practice degree at The Central School of Speech and Drama has an interesting method by which students establish assessment criteria. Following the initial establishment of peer assessment criteria devised by the whole class (described as the 'ground rules'), each student group was given the opportunity to agree two additional criteria, in order to enable each group 'to reflect aspects of individual contribution to group performance'. The intention of this approach was one of encouraging student ownership and

understanding of the assessment system, allowing students to be assessed as proto-practitioners, team members and individuals within the same format.

The tutor reported that a few students expressed a concern that some staff were reneging on their role as judges and experts, but a far greater number appreciated the comparison and sharing of practice that the assessment of other's work afforded them. The tutor believed that it was problematic for some students to be objective and not to assess others based on friendship or enmity.

It was considered that the benefits far outweighed these possible shortcomings: for example within group and process skills it was possible for students to discover that a person might not have been the best or most productive practitioner, but that their drive and enthusiasm within the group made them a first choice group member.

The basis on which students developed their peer assessment criteria has relevance for this thesis. Particularly interesting is the notion of describing criteria in terms of 'drive and enthusiasm' - personal qualities or attributes that indeed can make a great difference to group dynamics and the achievements of the members as a whole. Not entirely clear from the report, however, is the chronology of the criteria development. At what stage did students specify the additional criteria? If these were added prior to the assignment activity, then did the group anticipate that particular students would demonstrate qualities of 'drive and enthusiasm'? If they were been added at the conclusion of the assignment then the kind of qualities and attributes that its individuals demonstrated, would have been clear to all. If so, are peer assessment criteria, which are based upon these qualities and added post-event, credible? The report, does not, unfortunately address such questions. This thesis does, in its exploration of peer assessment criteria arising from previous band rehearsing experiences, as discussed particularly in Chapter Six.

1.3.6 Group characters: shooting stars and free riders

This section considers individuals and their particular characteristics which may become manifest during group work. One would expect to see a range of personalities, interpersonal characteristics and attributes of individuals in any group. For example, in

the context of band rehearsing, there may be students who appear shy, domineering, relaxed, earnest, introverted or self-centred. Two types of individuals who may be considered as being demonstrably different from the majority, due to their appearance at the extremities of group work effort and achievement, are known in the literature as the 'shooting star' and the 'free-rider'.

The 'shooting star' describes a high achieving individual in group work, both in personal performance as well as in their interpersonal skills. The lofty award of a 'shooting star' probably ought not to be one that is conferred solely on the basis of being the highest rated individual in a group assessment. If it were awarded only on that basis, then almost every band would contain a 'shooting star' and so would devalue its status. Perhaps it ought to be reserved for the highest rated student in the group who is also the best by some distance from the runner-up.

The term 'free-riders' typifies lazy individuals who, through their lack of effort, poor attendance, or contentment with letting the others do all the work, handicap the group. Other terms used in the literature that have the same meaning include 'freeloaders' (Nicholson & Ellis, 2000) and 'freebooters' (Doran *et al.*, 2000). Similarly as for shooting stars, it may be that the lowest-rated student in a group assessment by some distance deserves to be described as a free-rider. It is also possible, of course, that although effort was forthcoming, the student was demonstrably extremely poor in other areas. If so, then applying the label of free-rider may be unfair to a student who, despite their best endeavours, is simply very weak. Tutor knowledge of an individuals' effort, ability and understanding of the subject area is a key consideration also in deciding the appropriateness of using such labels. This thesis adopts these loose definitions for the top-rated shooting star and the lazy, low-rated free-rider.

Another characteristic of both shooting stars and free riders is one that is related to self-assessment. Lejk & Wyvill (2001a), for example, suggest that shooting stars often under-estimate themselves in a self-assessment activity, compared with how their peers rate them. The reverse appears to be the case for free riders: they typically over-estimate their achievements compared with how their peers assess them. This thesis, particularly in chapter six, extends our understanding of such characteristics and provides new knowledge about shooting stars and free riders in peer assessed group work.

Free riders and shooting stars have relevance for this research as both characteristics can be recognised and found among various student bands during their rehearsing and performing. Peer assessment may be helpful, not only in the identification of shooting stars, for example, but also in understanding more about their particular qualities and attributes evident in rehearsing and performing, in order to learn something about how to assist others in their aspirations of becoming shooting stars themselves.

Similarly, knowing a little more about how students respond to free-riders through peer assessment may help tutors in developing their group work practice and free-riders themselves, in learning more about their personal qualities and which ones peers have indicated need improvement.

Summary

The two concluding themes of this section on group work and peer assessment have been concerned with the *Assessing Group Practice Project* and the issues related to shooting stars and free riders. Four case studies reported in the *Assessing Group Practice Project* were particularly relevant to peer assessment and this enquiry. The example from the University of Ulster illustrated a resistance by students to tutor-intervention in the formation of group membership and also the use of peer assessment for grading purposes. The second example, from the University of Leeds, describes another basis for group formation and raises questions about peer-to-peer feedback that is accessed by the tutor. The third example, from Middlesex University, illustrates the importance of a distinctive context through which students' personal qualities and attributes become significant. The final example, from the Central School of Speech and Drama, describes an approach in which assessment criteria are developed during the assignment, rather than being prescribed in advance. In the literature on group work and peer assessment, reference is sometimes made to the problem of the free rider. The typical characteristics of free riders, together with those of its antithesis, the shooting star, complete this section on group work and peer assessment.

1.4 The Process of Peer Assessment

As with any method of assessment, a peer assessment system involves an operation of particular process features and mechanisms. The 'process' or 'operational mechanics' (Brown, 1999) of peer assessment are terms that are used here to describe the procedures that are applied over the duration of a peer-assessed activity. Typical examples of process features would include the establishment of assessment criteria, the marking system and the method of feedback. These features, which appear in many modes of assessment, are also considered fundamental processes of peer assessment. Indeed, assessment criteria and feedback has already been the subject of some discussion in this chapter.

The following survey of the literature, however, suggests that assessment criteria, marking systems and feedback methods within peer assessment can be distinctive and often quite complex processes.

There are, of course, many process features that can impact upon the peer assessment. For example, there may be decisions to be taken regarding the sequence, timing or emphasis of particular process features, especially where these may involve an assessment of a creative activity (e.g. band rehearsals) together with an assessment of the creative product (e.g. a performance). Other process features may present alternatives that require a decision: should the peer assessment of a group-based activity by its group members be conducted in private or through collaborative group agreement?

These, then, are examples of some of the features to be found in the 'process' of peer assessment and which will be examined in the literature. In order to discuss the key issues, the survey is structured into the relevant themes: intra-peer and inter-peer assessment; establishing peer assessment criteria; Orsmond, Merry & Reiling's (1996, 1997, 2000) research; category-based or holistic criteria; Leijk and Wyvill's research (2001a; 2001b; 2002); how many criteria; feedback and peer assessment; marking issues.

1.4.1 Intra-peer and inter-peer Assessment

This section, despite its brevity, has importance for defining and explaining the nature of the peer assessment relationship. Peer assessment generally falls into two types, which are explained by Brown (1999) and contextualised in the curriculum area of popular music by Pulman (2004) and Pickford & Brown (2006) as *intra-peer assessment* and *inter-peer assessment.*

When students assess other students (for example, when band performances are assessed by other students) this is referred to as *inter-peer assessment.* Alternatively, when students assess other students with whom they have been working (for example, when a band member assesses the contribution to rehearsals of fellow band members), this is known as *intra-peer assessment.* Similarly, as explained at the outset of this chapter, the term 'assessment' is used generically; *intra-peer assessment*, for example, embraces the activities of intra-peer support, intra-peer feedback, and intra-peer learning.

As these two types of peer assessment are referred to extensively throughout this thesis, it is important, therefore, that their meaning is defined at this point. The next section considers the literature on peer assessment criteria.

1.4.2 Establishing peer assessment criteria: who is responsible, who has ownership, and what are the opportunities for learning?

Along with the very considerable attention that has been given to assessment criteria generally, there is also a significant body of literature devoted to peer assessment criteria in Higher Education. Establishing criteria for students to use in the assessment of their peers is an area that reveals an array of issues and opinions in the literature. Procedures for generating assessment criteria are proving to be more complex than appears at first sight, and a number of key questions are raised. For example, *who* should be responsible for devising peer assessment criteria - students, tutors, or student-tutor collaboration? How *many* assessment criteria should be employed and *why*? Should the criteria be founded upon a category-weighted basis or be developed instead along holistic approaches?

A helpful definition of the term 'criterion' with regard to this section, is that of Sadler (1978): 'a distinguished property or characteristic of any thing, by which its quality can be judged or estimated, or by which a decision or classification may be made', (p.192).

A starting point in literature on peer assessment criteria is one that examines the origination of the assessment criteria itself. For peer assessment there is a body of opinion (Habeshaw *et al.*, 1993; Brown & Knight, 1994, for example) in favour of criteria that are determined by the students themselves since it is they, of course, who will be using such criteria to make their judgements. Importantly, student generated peer assessment criteria are advocated by many on the basis of the desirability of involving students in the assessment process in order to promote greater transparency. In the early and influential study of Goldfinch and Raeside (1990), conducted at a time when peer assessment in HE was in its infancy, these authors, although not employing student-generated criteria as was typical of the period, nevertheless acknowledged its desirability. In their research they suggest that, on reflection 'a discussion with the students could have helped to make them feel more involved and helped them to notice the skills as they were displayed in the group' (p.210). For example in 'Peer Assessment in Practice', a key text in the literature, Race (1998) provides helpful ideas for tutors to facilitate student-generated criteria. It is also important that students agree to the criteria and Gibbs (1999) suggests the practical step of disseminating copies of the criteria to those involved as soon as possible.

There are arguments in favour of involving students in developing their own assessment criteria. One of these, according to Prins *et al.* (2005), resides in the potential of peer assessment to 'thrive on interaction' and so provide a natural setting through which students can determine and negotiate assessment criteria. Importantly, Boud & Falchikov (2006) emphasise the wasting of learning opportunities that are available here, through omitting the involvement of students in establishing criteria in these key stages. This contrasts somewhat with the mild endorsement offered by O'Donnell and Topping (1998) that students' involvement in establishing peer assessment criteria is 'desirable'. However, one of the most significant peer assessment projects in the literature, conducted by Leijk and Wyvill (2001a), used tutor-imposed criteria. Butcher (2004) depicts the act of the tutor relinquishment of control over criteria as a gradual

process, described sequentially as 'yours; yours and theirs; theirs and yours; theirs' (p.4).

Not all studies support the student-generated criteria, however. It was found in the research conducted by Orsmond, Merry & Reiling (1996, 1997, 2000) that students were less able to discriminate between assessment criteria they had constructed themselves compared to tutor-imposed criteria. Also, it was found that students' own criteria did not enhance student-tutor or student-student marking agreement. Their research is an important contribution to the debate of peer assessment criteria and will be considered separately in the next section.

Some studies raise concerns about the quality of student-generated criteria, in terms of their clarity and meaning. The guidance that was offered by Pond & ul Haq (1998) nearly two decades ago was that criteria should be clear and category-weighted (i.e. each criterion attracting a separate weighted mark, as explained later in Section 1.4.2.2). They use the term category-weighted to mean that the assessment should comprise a number of discrete criteria, each of which having a specific percentage weighting of the whole. Mindham (1998) draws attention to the necessity, in her view, that such criteria should be written in a plain and readily understood language. It is also a view that is shared by Doran *et al.* (2000), who emphasise that 'consistent, detailed and explicit criteria for assessment are fully discussed and agreed by tutors' (p.70). Doran *et al.* argue that the use of explicit peer assessment criteria also limits the degree of subjectivity and impression marking. Similarly, Purchase (2000) describes using 'firm objective criteria' as 'safe' (p.341).

Rust *et al.* (2003) in their study of business students and Bloxham & West's (2004) account of peer assessment in sports studies, have each investigated the problems associated with how students can become absorbed into the assessment culture of their disciplines, with recognition that certain written criteria may not be enough to make clear their meaning and intentions to novice students. A problem is that criteria tend to be articulated in a written format only and this medium may not always be sufficient in conveying subtle meanings of a non-discursive type. Recognising this, Rust and his colleagues considered the problems of initiating students into assessment culture and the need to develop a sense of 'connoisseurship' that is dependent upon communication and

experience in a wider sense 'involving observation, imitation, dialogue and practice' (p.152). Such acquired knowledge they describe as 'tacit', and the 'tacit' nature of assessment criteria is a real problem because of the difficulty of transferring such understanding to others. Socialisation, Rust suggests, is required to assist in a transfer of 'tacit' knowledge (p.161). This issue may perhaps, be likened to certain aspects of situated learning theory. Rømer (2002) for example, in his account of the situated learning theories of Lave & Wenger (1991) in relation to assessment, draws attention to the concept of 'the criteria of the assessors' being found in 'a community of knowers' (p.235). Such criteria, Rømer argues are tacit and intuitively grasped and, according to Morgan & Wyatt-Smith (2000), this knowledge 'exists largely in unarticulated form as *lore*' (p.134).

Bloxham & West's (2004) small-scale, largely qualitative study, involving sports studies students using assessment criteria to mark poster displays by their peers, investigated the use of a number of college-wide assessment criteria. The study, which included a group interview and questionnaires, specifically explored how systematically teaching about the assessment might improve students' understanding of assessment criteria in preparing their posters and peer marking these. The writers report some success in that students 'had engaged with the criteria in preparing and marking the posters that a high proportion were able to accurately predict the grade they had achieved.'(p.731).

Similarly, Langan *et al.* (2005) state that in their study a firm understanding of the assessment criteria was associated with greater validity. Although few would disagree with the need for students to have sufficient comprehension of the meanings of assessment criteria, this does not necessarily equate with criteria that are presented in an explicit manner.

Boud & Falchikov (2006) make the point that assessment criteria are rarely specified in an explicit real-world context. Norton (2004) develops a counter argument questioning the appropriateness of explicit criteria because of the danger of rigidity or inflexibility, which could limit students to focus on the purely visible. Indeed by adopting explicit and detailed criteria intentionally, Norton argues that this may have a 'deleterious effect' upon the peer assessment process. She suggests that we should come to replace the term

'assessment criteria' with 'learning criteria' in order that the focus is not on the purely visible, but includes wider demonstrations of learning that are more centred on the individual. Might there be peer assessment contexts where by the nature of the activity – group rehearsing situations extending over a sequence of rehearsals, for example – Norton's concept of assessment using learning criteria is more appropriate? If learning criteria should be so centred on the individual, then what are the applications and implications for individual members of a band who are working together in their rehearsals and performances? These interesting and thought provoking questions as a consequence of Norton's 'learning criteria' are revisited in Chapter Three.

Ballantyne *et al.* (2002) report that while the process of student-generated criteria was successful with a smaller group of students it became 'too unwieldy for use with large classes' (p.432) when they attempted to develop similar criteria across several tutor groups. When students devise their own criteria, as well as acquiring 'ownership', there may be a sense in which, through the generating process, their criteria become 'intimate' to themselves. If so, would attempts to extend such notions of intimacy across large numbers of tutor groups become problematic, if not contradictory?

Finally, we are reminded by Hanrahan & Isaacs (2001), Smith *et al.* (2002), Ballantyne *et al.* (2002), Pope (2005) and others, that not only is the setting of clear peer assessment criteria important, but equally so is training students in their use.

Overall, a conclusion may be drawn from the published literature on peer assessment criteria from 2000 onwards, that it is the norm to involve students in developing their own assessment criteria. However, this is an area that contains a number of pitfalls and remains problematic as noted in the investigations by Orsmond and Merry (2000). Their research challenges the arguments of those in favour of involving students peer marking using their own assessment criteria. It is an important contribution to the literature and necessitates further discussion.

1.4.2.1 Orsmond, Merry & Reiling's research: the problem of students' comprehension of assessment criteria

Three studies carried out by Orsmond, Merry & Reiling (1996, 1997, 2000) investigate the extent to which first year undergraduate biology students understood the meanings of the marking criteria that were employed. The criteria were imposed by the tutors and comprised eight equally weighted categories. Verbal and written briefings were given before the assessments and confirmation that students understood the terminology was obtained. The two earlier studies (1996, 1997) compared tutor marks and student marks using the given criteria. The authors concluded that students had a different understanding of some of the individual criteria compared to that of their peers and tutors. If so, this would have a serious impact upon the effectiveness of peer marking using a set of criteria that was not adequately understood by the markers themselves.

Their later study (2000) involved students constructing their own criteria to judge a poster presentation in discussion with a tutor. An evaluative questionnaire was given to the participating students at the end of the activity. They concluded, as described earlier, that students might be less able to discriminate between assessment criteria they have constructed themselves compared to tutor-imposed criteria and that students' own criteria did not lead to improved student-tutor or student-student marking agreements. Their study has had some influence in the literature; it indicates caution, and presents a challenge to the arguments of those who are in favour of involving students in developing assessment criteria. The writers suggest, agreeing with Brown & Knight (1995) and Habeshaw *et al.* (1993), that 'a possible way forward could be joint tutor/student construction of the marking criteria' (p.30). Yet in a subsequent article Orsmond *et al.* (2002) appear to contradict this suggestion by stating that a' joint construction of marking criteria did not identify a method of using peer and self-assessment that achieved greater understanding of the marking criteria by students.' (p.310). This later research also explored the use of exemplars in order to help students acquire a better understanding of assessment criteria and subject standards. Their findings confirmed that exemplars did indeed help students to demonstrate a greater understanding of marking criteria and that it formed a focus for meaningful formative feedback and more objective judgements.

Their study raises questions, however, about the chosen method of participation. Despite involving every student within each group in a discussion about the meaning of their criteria, the peer assessment itself was conducted by pairs of students only, rather than by whole groups. Why this method was chosen is not explained. As the criteria were created and 'owned' by the group, it is surprising that an intra-group assessment was not employed in order to include the input of all the group participants who shared in developing the criteria, rather than pairs of students.

An encouraging outcome however, was that through the development of the criteria students themselves acquired a kind of special understanding; the researchers offer an interesting and speculative thought:

> they may have developed a sense of "ownership" which related to both the *meaning* of the criteria and the *worth* of the criteria in terms of marks to be awarded. Students may feel that because they constructed the criteria they are arbiters of the quality of the criteria expression and/or of the subject matter the criteria expresses (p.320).

Although in their research, students appeared to be less discriminating about their own criteria compared with those of tutors, they may have begun to comprehend the meanings of the criteria within their shared student community. If so, might this kind of understanding be likened to the aforementioned espousal of 'tacit knowledge' (Rust *et al.*, 2003)?

1.4.2.2 Category-based or holistic criteria?

There is also some debate in the literature concerning the purpose of peer assessment criteria that are based either upon a number of categories, or, that of a single holistic descriptor. To recap, category-based criteria typically refer to organised, discrete categories or dimensions, each of which may be accompanied by a percentage weighting to each criterion (category-weighted). Holistic criteria are usually characterised by having either a single overall dimension or general theme in which there is no attached weighting. The term *analytic* is sometimes used as an alternative to *category* especially in the United States. Rust *et al.*, (2003) and Biggs (2003) also provide a helpful account and survey into the construction of category and holistic assessment criteria.

Studies that specifically address this aspect in the context of peer assessment, however, are less common than those that relate across assessment in general. Nevertheless, an appreciation of the rationales behind the establishing of each and the practicalities of using such criteria in order to judge peers' work is important when contemplating peer-assessed activities.

In some often-cited earlier studies, there appeared to be a prevalence of category-weighted criteria (Pond & ul-Haq, 1997; Goldfinch & Raeside, 1990; Li, 2001). The research by Leijk & Wyvill (2001a; 2001b; 2002) has been referred to earlier in this chapter during the discussion on the generating of criteria. Their work specifically investigated aspects of category-weighted and holistic criteria and forms a useful contribution to the area. These articles remain a useful source for research into this area and are of some relevance also to this investigation; they are considered in the following section.

1.4.2.3 Leijk and Wyvill's research investigating holistic and category-based approaches

Their papers present the results of an experiment, which compared two peer assessment approaches with students studying a year two undergraduate module in Business Systems Analysis at the University of Sunderland. The first approach adopted an holistic criterion, where each student awarded just one grade to each of the other group members, based on the word 'contribution'. The second approach employed a category-based assessment employing six tutor-imposed criteria comprising:

Motivation, Responsibility/Time Management
Adaptability
Creativity/Originality
Communication Skills
General Team Skills
Technical Skills

Tutor-imposed criteria were employed in order 'that all groups should use the same set of categories for the category-based approach.'(p.63). Why they did not consider,

however, involving all the students in deciding the criteria and agreeing upon its use for all their groups was not explained.

An important finding overall in their research raises questions about the use of category-based assessment methods. They investigated measures of intra-group marking agreement, finding that although considerable agreement occurred in both approaches, the holistic method displayed the strongest agreements. In their research, holistic assessment criteria led to a much higher proportion of groups who awarded each other equal marks, and Leijk and Wyvill concluded that it led more directly to a measure of each member's contribution to the group effort. This latter point was also supported in their survey of the participating students. It could also be argued, however, that although a high proportion of students awarded each other equal marks using the holistic method, this produced a less differentiated outcome.

That there was a greater consistency of agreement reported among students when using an holistic criterion is perhaps not surprising. Sharp (2006) explains that whilst students may have the tendency to award an average mark for a single holistic judgement, it is statistically less likely that they would award the same mark across a number of categories. Although a student may have a high or low opinion of another student using one criterion, it is less likely that the student would maintain a uniform opinion across a number of categories and 'the use of a multiplicity of criteria or categories will probably lead to regression to the mean' (p.331).

Important questions arise from the choice of categories and the nature of the assignment. It is implied that the categories are equally important, yet, because of the nature of the group task, would it have been possible that all group members would have had the opportunity to have participated or demonstrated their achievements across the six categories? If the nature of the assignment related exclusively to team working activities then it would not be unreasonable to have categories relating to group working skills. However, depending on the nature of the assignment, most groups would, quite reasonably, allocate specific tasks to its individual members at some stage. Distributing the workload among group members in this way, however, may render some of the categories as inapplicable or inappropriate for certain members of their group. If peer assessment is concerned primarily with obtaining a measure of each member's

contribution, then category-based criteria, they argue, are in some conflict with this summative purpose.

Although Leik *et al.* raised questions relating to the nature of the group work activity and the relationship of the category-based criteria to individual students, they did not pursue these in their paper. These are at least two key issues arising from their work that have a bearing on this thesis.

Firstly, ought criteria to arise from the nature of the group activity or from the summative requirements specified in the assignment brief? Imposing a set of criteria based upon the latter but without reference to the nature of the activity might be dangerous and could lead to a distorted assessment of what really happened. Understanding the assessment activity and its setting are significant indicators and a prelude to the establishment of appropriate category-based criteria.

Secondly, ought consideration be given towards the particular and individual relationship of students to the assessment criteria? By this, should we be asking to what extent can students individually as well as a whole relate to these criteria? Are the peer assessment criteria distant, impersonal and generic or student-centred, personal and individually specific? As noted earlier in this section, Norton's concept of learning criteria that relate more to individuals becomes particularly relevant in addressing such questions.

Leijk & Wyvill (2002) also conducted a survey among the students who participated in their experiment, as an attempt to gauge students' attitudes to group assessment using a *Likert* scale. The findings from this also supported the authors' earlier quantitative research that holistic assessment is "more in tune" with the purposes of group assessment than category-based methods. Their findings also raise the question of whether the method of assessment used (i.e. using category-weighted or holistic criteria) has some influence upon how students work with each other in their groups.

Category-weighted criteria are more appropriate, they suggest, for formative assessment, a view echoed by Bushell (2006). Although their assertion is not explained in particular detail, the implication is, perhaps, that setting category-weighted criteria at a formative

peer assessment stage, encourages learners to focus and apportion their efforts into specific, tangible and named areas. One of these would almost certainly include the development of individual skills or particular abilities as a kind of pre-requisite to comprehending the 'big picture'. Adopting such criteria might, as Goldfinch & Raeside (1990) suggest, allow learners to notice such skills being developed within their group members.

Ultimately however, it is holistic criteria that are regarded by Leijk & Wyvill as more appropriate overall for summative and group-based assessment, where presenting a finished group artifact for marking purposes is of superior importance to a focus upon individual parts.

From this survey of a somewhat specialised area of the literature, it is possible to conclude that holistic criteria are often superior to category–weighted criteria in peer assessment, although the latter have an important role in formative assessment. Nevertheless, these issues are far from being clear and other important factors may include the nature of the activity to be peer assessed and how the relationship of individual students to the criteria is specified.

1.4.2.4 How many criteria?

An interesting question to ask would be one that ascertains the number of criteria that we are asking students to generate. Using the minimum of just one criterion frequently leads to descriptors such as 'contribution to the group'. The distinction between holistic and category types of criteria can, however, become blurred. Although Gupta (2004) describes using a 'contribution' criterion, she explains that it is based upon 'participation', 'effort' and 'sense of responsibility'. It is not clear, however, by asking her students to focus on those three elements, whether she is employing three separate categories rather than an all embracing holistic like 'contribution' criterion.

It is a concern of many students that recognising their own individual effort to a group project be reflected in the assessment stage (Mindham, 1998), desiring that their efforts may be identified in the peer assessment criteria. The holistic 'contribution' criterion is a simple solution and popular method for judging an individual's contribution to group

work. One purpose for using a number of criteria, however, is that of obtaining a range of assessment information across various dimensions that represent the individual's contribution.

The study by Langan *et al.* (2005) involving peer assessment of oral presentations on environmental or biological topics, describes that although there were 41 students participating, the tutors decided that 12 randomly selected students only, should be involved in the development of the assessment criteria. This they justified, somewhat unconvincingly, on grounds of balance between gaining many opinions and managing a meeting whereby all in that meeting could participate 'without being crowded out' (p.24). Students were supplied with subject benchmark statements with which to 'guide' the development of their assessment criteria. Three category-weighted criteria were established: Content (40%), Presentation (40%) and Structure (20%). Each of these criteria had two sets of accompanying descriptors, one relating to a 'threshold' achievement and the other relating to 'excellence'.

Although it was stated that 'all learners were encouraged to take ownership of the development of their projects' (p.24) this desire would seem to be clearly at odds with how the students were directed to devise their assessment criteria. Somewhat disappointingly, those students who participated in devising the assessment criteria did not achieve higher grades for their presentations.

Goldfinch & Raeside (1990) present three examples of category-criteria that they have used in mathematics. The first comprised a total of seven categories (equally weighted) for their group mathematics peer assessed project of 1989:

Degree of participation
Leadership of group
Suggestor of ideas
Ability in understanding what was required
Ability in doing the analysis/computation
Consolidating and interpreting results
Sorting out problems

Another example included in their research comprised six criteria, including 'coming up with something useful', a somewhat vague criterion and open to interpretation in a number of ways. A final example comprised eight categories including 'extracting something useful from the ideas', 'sorting out problems' and 'keeping the group going, particularly in difficult patches'. All categories were equally weighted and used a 0-4 marking scale. Orsmond *et al.* (2002) used a joint tutor-student construction of category-criteria in their study. Six groups of students were involved and each group created four criteria. The authors do not explain if the groups were required to produce four criteria or whether this was a remarkable coincidence.

Tutors may need to consider the implications of managing peer assessment over a large number of category criteria. For example, asking students to discriminate in increasing detail across a large number of categories can provide a useful learning experience as well as yielding a greater range of information. This potential benefit, however, ought to be weighed against the capacity of the students to be properly equipped and able to do such work, and the appropriateness of employing a large number of marking categories. Many reported peer assessment studies with a category-based approach employ between three and eight categories. Using large numbers of categories is rarely found in the literature and this might suggest that tutors have anxieties of their students' capacity to manage the process and discriminate appropriately. Alternately, restricting category-based assessment to just two categories, might raise the question of whether using a single holistic criterion would be better suited to the purposes of the assessment.

Miller (2003) describes a study involving his physical therapy class. In particular he investigated the use of highly specific assessment criteria. The course followed a problem-based format involving groups of around five students culminating in a product-type of peer assessment: a group oral presentation. For the first cohort, a total of five tutor-imposed category criteria were employed, using a 0-4 rating scale for each criterion. The next iteration of the course however, employed a total of 25 categories with each criterion marked across a 0-4 rating scale as before. Miller reported some positive findings as a result of using more criteria specificity. For example, peer feedback was consequentially greater, yielding a wider range of scores and with so many items, students 'felt more comfortable about downgrading certain areas of performance as it would not have a large impact on the overall score.'(p.383). Miller rightly asks that

educators need to consider whether employing such a range of specific criteria might be at the expense of qualitative feedback. However, he does not discuss whether asking students to use this large number of specific category criteria might alter their perceptions of the assessment. Sluijmans *et al.* (2002) make a similar point that an increase in the number of criteria is not necessarily accompanied by adequate and correct usage.

1.4.3 Feedback and peer assessment: what are the issues?

Den Berg *et al.* (2006) argue that peers learn more when the peer assessment process incorporates peer feedback. Questionnaires used as a part of their research suggested that the students valued written and oral feedback in 83% and 78% of returns respectively. In the paper by Bloxham & West (2004) that was discussed earlier, their students were invited to describe the main points of the feedback they received. The directness of feedback from student to student was interpreted positively:

> The straightforward way by which the students seemed to understand and articulate their feedback may be a positive result of receiving feedback from other students expressed more plainly than tutors might be able to do. However even though students claimed to understand the feedback, it was not something with which they always agreed. (p.721).

Although training was given to students regarding the assessment criteria and marking, there was no indication of whether students had been similarly prepared for the giving or receiving of feedback.

The provision of regular and ongoing feedback that Bransford *et al.* (2000) maintain, is an important part of the process that enables students to improve or revise their work. Similarly, Drew (2001) explores students' perceptions of the link between assessment feedback and learning. Topping *et al.* (2000) similarly agree and refer to 'rich and detailed qualitative feedback information about strengths and weaknesses' (p.150). Rust *et al.* (2005), Price & O'Donovan (2006) and Price *et al.* (2007) require, in their social constructivist assessment process model (considered elsewhere in this chapter), that students actively engage with feedback. Race (1999) argues that peer assessment feedback should be meaningful and helpful and exhibit increasing 'richness'. Devising training activities in order to develop feedback that is both 'meaningful' and 'helpful'

might help improve the language in which students communicate their feedback to one another. An interesting training activity that was used by McDowell & Samball (2005) in a peer assessment assignment undertaken as part of an Early Childhood Studies year one course was designed to help individuals formulate feedback to fictitious students.

If feedback from peer assessors is intended primarily for the fellow peer recipient, then this raises questions about tutors' involvement in such transactions. To what extent should tutors be privileged in accessing peer-to-peer feedback? If tutors' inclusion in peer-to-peer feedback is a necessary part of peer assessment training or justified on grounds of enhancing the transparency of the process, this should be explained to those involved. The role of the tutor, the respect for students' privacy and the objectives of providing feedback in the peer assessment are all factors in deciding the appropriate feedback mechanism. For O'Donnell & Topping (1998) the exchange of peer feedback can be a particularly challenging moment in the peer assessment process. Where detailed feedback is anticipated, then they recommend that this is not accessed by authority figures.

1.4.4 Marking issues in peer assessment

The literature relating to peer marking and quantitative analysis in peer assessment is particularly extensive, reflecting the large number of published fieldwork studies that are now available. The purpose of this section is to survey the key issues and themes that appear to emerge from a study of this literature. As such, there are perhaps three areas that appear to attract attention in the literature: the particular marking process employed (including the so-called 'zero-sum' method), the chosen rating scale and range, and quantitative analysis in studies.

The first issue concerns the various marking methods adopted by tutors in peer assessment. In inter-peer assessment, peer markers may only effectively ascribe a mark for a product (a performance, for example) created by an individual or by a group. Intra-peer assessment, however, can be used in order to determine marks based upon the contribution of individuals to a group project. The process by which individuals are allocated their marks requires tutors to decide on an appropriate marking method. Sharp (2006) and Gatfield (1999) provide a brief discussion on the development of various

approaches to this. Lejk *et al.* (1996) describe a minimalist method consisting of the group members receiving the same grade unless one is felt by all the others to be a free-rider, in which case the tutor decides whether to moderate the free-rider's mark. Another method, and one that is perhaps the most widely employed, consists of inviting each group member to apportion 100% among the group members. The percentages allocated to each individual by their peers are totalled and a further calculation, involving the multiplication of the 'group mark', is performed. In practice, students' final mark is obtained from the totalled intra-peer marks, dividing each by their mean, and multiplying them by the 'group mark'. It can be expressed thus:

$$\frac{\text{Totalled intra-peer assessment mark of individual}}{\text{Mean totalled intra-peer assessment mark of group}} \times \text{Mark awarded to the group}$$

This process is sometimes described as the 'zero-sum' method (Sharp, 2006) because any student who is peer assessed as providing zero contribution receives zero marks.

Sharp (2006) takes issue with this method, describing it as 'abdicatory', accusing tutors of 'turning the whole lot over to the students to divide up as they see fit' (p.330). Bushell (2006) and Lejk & Wyvill (2001b) explain particular discrepancies and bias that may occur in peer marking methods. They note a tendency for able students to under-rate themselves and poorer students to over-rate themselves.

Li (2001) also notes various anomalies in peer marking methods, and proposes a solution, entitled 'normalisation' which, he argues, makes corrections for students attempting to inflate their own marks by awarding inappropriately low marks to their peers. Pioneering methods explored in earlier studies are often based on a combination of approaches such as those of Goldfinch & Raeside (1990), Conway (1993) and Raqif & Fullerton (1996). Gatfield (1999), for example, describes a marking method in which 50% of students' final grade comprises a mark for their product with the remaining 50% based upon a 'zero-sum' type distribution.

The various rating scales that are employed in peer assessment can also be an issue. A significant number of studies adopt 5-point scales (for example, Oldfield &

McAlpine, 1995; Orsmond, Merry & Reiling, 1996; Gupta, 2004; Johnson & Miles, 2004; Den Berg *et al.*, 2006) and their reasons in so doing are usually based upon the view that deciding marks out of five allows a reasonable discrimination to take place without being operationally unwieldy. Rating scales of less than 5 are uncommon in the literature, although larger scales are occasionally employed (for example, the 7-point scale adopted by Ballantyne, 2002). Miller (2003) cites a number of studies employing 4- or 5-point scales and draws attention to the narrow range or bunching of marks at the top end that, he suggests, typically occurs in peer assessment when adopting these rating scales. In an attempt to combat this tendency, as mentioned in Section 1.4.3, he introduced a marking instrument comprising 25 tutor-imposed criteria, each based on a rating scale of 0 (unsatisfactory) – 4 (excellent). This provided for a possible total of 100 marks, which produced lower mean scores. Although it may have provided more quantitative discrimination this appeared to be at the expense of qualitative feedback. Previously, when using 5 items, students provided a comment also, but this was deemed to be impracticable across 25 items. Lopez-Real & Chan (1999) and Oldfield & MacAlpine (1995) however, share the view of several writers in advocating simplicity of peer assessment apparatus.

Determining which range to use for the marking scale is another factor. The most common range used appears to be 1-5 (for example Pond & ul-Haq, 1997; Cheng & Warren, 1999; Sullivan *et al.* 1999; Li, 2001). Some prefer the use of a zero to record zero effort and adopt a 0-4 range (Orsmond *et al.*, 2000). However Lejk & Wyvill (2002) adopted a marking range based upon that of Goldfinch (1994) comprising: -1,0,1,2 and 3, (–1 being described as a hindrance to the group).

The literature of quantitative analysis in peer assessment studies is extensive. Many studies feature statistical tests involving correlation analysis, analysis of mark differences and of variance. These often involve comparisons between peer marking and tutor marking, and between self- and peer-assessment. These aspects are of particular interest to those who wish to see evidence that student peer marking can be consistent and dependable enough to be used for summative purposes. Magin (1992; 1993; 1996; 1997; 2001) has published much valuable work in this field of the literature.

Summary

The literature on process features surveyed a number of key issues involving the operational mechanics of peer assessment. A consideration of the various process mechanics within a peer assessment system involves issues that can be difficult and complex. This section commenced with the concept of intra-peer and inter-peer assessment and defining who is involved in the process. The literature on peer assessment criteria revealed the extent of the many intricacies that are involved in this area, for example, involving students in establishing the criteria and the consequences in so doing. One issue to emerge was that of students' comprehension of the criteria. There were concerns about peers' tendency to focus upon the purely visible and their need to acquire 'tacit' understanding of assessment criteria. Partly related to this was the involvement of wider demonstrations of learning that are more centred on the individual: Lin Norton's concept of 'learning criteria'. A lengthy survey was conducted on the issues of using holistic or category criteria. The literature on student and peer-to-peer feedback was examined which raised a number of questions concerning methods of giving and receiving feedback. Finally, this section presented an overview of quantitative marking processes employed in peer assessment. An important part of the process of peer assessment involved methods for identifying and determining marks based upon the contribution of individuals to a group project. Issues concerning particular rating scales and ranges were considered and this section concluded with the observation that statistical tests, particularly analysis of variance and correlation analysis, were a feature of many peer assessment studies.

This chapter has so far surveyed the key themes and issues arising in the literature that may apply to peer assessment across the HE curriculum in general: peer assessment and learning; group work and peer assessment; the process of peer assessment. The next section revisits many of these issues in examining the literature of peer assessment in the curriculum context of music.

1.5 The Literature on Peer Assessment in the Music Curriculum

This section of the chapter is devoted to a review of the literature on peer assessment specifically in HE music curriculum settings. In comparison with other curricular activities (for example disciplines involving peer assessment in a visual medium, such

as posters designs), music can be considered to be a little under-represented as measured by published peer assessment studies. Activities of group performance of music are even less common, as most reported peer-assessed performances employ a solo recital format, rather than one that is ensemble-based. Peer assessments involving popular music styles are somewhat rare. Studies that link the use of peer assessment with popular music groups are practically non-existent.

Significantly, it is in this latter assessment context – peer assessment and popular music group work - that the thesis is located. Another consideration is that of the HE curricular area of popular music. Whereas the study of western classical music enjoys a long tradition in academia, the undergraduate study of popular music is a relatively recent development. Just a handful of courses were available before and during the 1990s in UK institutions; it is from the new millennium only that the number of HE institutions offering popular music based programmes have significantly increased. This is a consideration that is relevant in exploring the extent of the available literature. Consequently, as a relative newcomer to HE, the available literature in terms of the learning, teaching and assessment of popular music at undergraduate level is not particularly extensive.

Fortunately, a welcome and valuable contribution to the literature in the UK was provided with the appearance of the *Peer Learning in Music* project published by University of Ulster in 2000. This initiative, supported by The Higher Education Funding Council for England and the Department of Education for Northern Ireland through the Fund for the Development of Teaching and Learning, remains the primary source for research into peer learning and peer assessment in music. This part of the literature review necessarily examines a number of studies drawn from the *Peer Learning in Music* project.

The following studies have been identified as key research literature, in which the tutors' experiences of peer assessment in music suggest a number of practical insights and theoretical principles that contribute to our understanding of the assessment process. These are categorised by HE institution, in a loose chronological order.

Kingston University:

Searby & Ewers (1996) Peer assessing composition in higher education

Searby & Ewers (1997a) An evaluation of the use of Peer Assessment in Higher Searby & Ewers (1997b) Education: a Case Study in the School of Music, Kingston University Peer Assessment in Music

Ulster University:

Hunter & Russ (1996) Peer Assessment in Performance Studies

Hunter (1999) Developing Peer Learning Programmes in Music: Group Presentations and Peer Assessment

Peer Learning in Music (2000)

James Cook University

Daniel (2004) Peer assessment in musical performance: the development, trial and evaluation of a methodology for the Australian tertiary environment

University of Western Sydney

Blom & Poole (2004) Peer assessment of tertiary music performance: opportunities for understanding performance assessment and performing through experience and self-reflection.

1.5.1 Kingston University

One of the earliest reported activities using peer assessment in the HE curriculum area of music was at Kingston University. Peer assessment was introduced into composition in 1992 and performance in 1994. A description and evaluation of these developments were subsequently published in three papers (Searby & Ewers, 1996, 1997a, 1997b). Their peer assessment activities in composition are also included as an exemplar in the *Peer Learning in Music* project (2000).

They originally introduced peer assessment into the second year composition course in order to reduce lecturer-marking time and to try to give speedier feedback to students. An extensive training scheme was introduced for students along with a set of generic guidelines relating to the administration of the process.

Assessment criteria were identified and negotiated by the students. They generated their assessment criteria through discussion that produced a number of lists, which were gradually refined and focused into a smaller number of general statements. The criteria for assessment were renegotiated each year; Searby & Ewers regarded the activity of devising criteria as perhaps the most important part of the process. However, they thought that it would be too unwieldy to ask students to additionally specify levels of achievement as they discovered the entire process was already very time consuming.

It was discovered that students preferred to assess in a small group, rather than as individuals. Generally, the student assessors' marks did not require extensive moderation and there was little evidence of over-generous peer marks being awarded.

Peer assessment was extended to performance during 1994. Before this is described, it is of interest to reflect on the order in which peer assessment was approached: composition to begin with and performance later. Many might consider that the peer assessment of composition is particularly problematic, due in part to the complexity of assessing compositions where the compositional process may be as critical as the product (for example Daniel, 2004). With performance, however, assessment is typically concerned only with summative assessment of the product - the recital or the gig. (That is not to say process aspects of performance, such as group rehearsing, practising and preparation, are less important; in assessment terms, however, they attract much less attention as either formative or summative assessment activities). Searby and Ewers state that 'peer assessment seems to work particularly well in subject areas where aesthetic judgements are important and knowledge and reasoning less so (e.g. composition and performance)'. Although many would agree in principle with this view (for example Brown & Knight, 1994), the activity of composition can, nevertheless, be dependent upon or involve an acquired 'knowledge' that is founded on theory and musicology.

For performance, students negotiated assessment criteria in a similar manner to that of composition. Groups or panels comprising 5-6 students assessed the performances, which were solo recitals. Marks were described as being 'negotiated' with a moderating tutor rather than the panel awarding a mark that is subsequently tutor-moderated. As such, this procedure might be described more accurately as a student-tutor collaborative

assessment, rather than purely student peer marking. The process appeared to work very well, except that some panels felt that tutors were ignoring their views, dictating rather than moderating.

Each of the activities involving composition and performance at Kingston can be described as inter-peer assessment of individual students. Group-work was not involved and the student assessors had not been directly working with the individual performers or composers.

In their paper evaluating their use of peer assessment, Searby & Ewers identify at least four different aspects in ensuring an effective process of peer assessment:

- the process of arriving at an appropriate mark;
- providing adequate written feedback;
- creating a speedy system for submitting, marking and returning work;
- ensuring students learn as much as possible from the process.

Later on, they suggest that students appear to work harder during peer assessment conditions, because it 'seems to provide a much more effective spur to the production of high quality work than does assessment by staff' (p.377). They also report that the responses from students in the second year of operation were more encouraging about the experience than those involved in the initial year.

In continuing to establish peer assessment, it appears, interestingly, that first year students exhibit lower levels of approval (although not in the case of music performance). This has necessitated the Music Department to issue a document on the thinking behind peer assessment, to be given to all students and lecturers at the start of each year. Moreover, they identified a need to integrate the process with visiting instrumental tutors who, despite being responsible for the individual instrumental lesson, were not presently involved in the assessment.

The real benefit of peer assessment, they concluded, was in the development of students' critical analysis. It also helped to make students 'much more critical and questioning in their approach to all their assessments, which in turn benefited both the students

themselves and the course as a whole' (Searby & Ewers, 1997b, p.7). This has given the tutors confidence to extend peer assessment to other modules: Music and Business; Music Technology; Performance and Communication, although it is not known whether the work in these areas has been published.

1.5.2 Ulster University

Taking place at a similar time to the peer assessment activities at Kingston University, were a number of initiatives at the University of Ulster (Hunter & Russ, 1996; Hunter, 1999) culminating in extensive contributions to the Peer Learning in Music Project (Hunter & Russ, 2000).

Peer learning and assessment commenced in 1992 (coincidentally at about the same time as Kingston University) on the BMus programme. The centrality of peer learning to the rationale and assessment process is evident throughout their work. Indeed, apart from the title of their first published paper 'Peer Assessment in Performance Studies', the authors use the term 'peer learning' or 'peer feedback' to entitle subsequent activities in other music modules. These include: Peer Learning in Baroque Studies and Renaissance Studies; Peer Learning in Trends in Late 19th and Early 20th-Century Music; Peer Assessment and Peer Feedback in the Teaching of Orchestration.

An initial impetus for introducing peer assessment of performance came from the students themselves. As tutors were aware that their students 'discussed their performances with their peers' they explored channels whereby 'students could, through acting as assessors, increase their understanding of what constitutes a good performance' (Hunter & Russ, 2000, p.23). An extensive training programme was introduced including performance seminars, 'shadowing' of senior peer assessors, seminars specifically designed towards the formulation of assessment criteria, mock assessments, review and analysis of previous video recorded peer discussions. They suggest that the preparation and training activities would require at least three hours in order to create the conditions for a successful peer assessment programme.

Year three students were formed into panels, usually of four, in order to assess second-year performances. The assessments were focussed upon 'product': solo performance

recitals. This process, for the assessment of individuals, was similar to that employed at Kingston University. The panels were selected by tutors upon the basis of providing an appropriate balance of instrumental expertise: ideally vocal, strings, wind and keyboard. The panellists regarded the collaborative assessment process as a very worthwhile experience, although it was reported that many were unhappy about having to award a mark.

The performance assessment criteria were renegotiated with each cohort and an elaborate process for generating these has evolved. Students generated non-weighted category criteria. Great care was also given towards asking students to define what they might expect in a first class performance, a lower second and thereafter the remaining classifications. This process went a stage beyond what was reported at Kingston University, where it was decided that asking students to provide level descriptors or classifications would be too unwieldy and time-consuming. It was clear that Ulster University considered that the time invested in this process was worthwhile in terms of assisting students' understanding of the criteria for achievement. Although not made explicit in their account, tutors, through engaging the group in developing assessment criteria, level descriptors, and reflecting upon their students' responses, would themselves have learned much from this experience. Indeed, it would have been interesting to learn a little more about the impact of the work at Ulster University for tutors themselves at an individual and personal level. For example, the impact would affect areas such as tutor-student trust, developing sufficient confidence to be able to manage the peer assessment process, attitudes to risk-taking, and relinquishing their traditional status as figures of authority and power.

Students raised some important issues, arising from the training sessions. Interestingly, these same questions, paraphrased below, frequently pose problems for tutors also.

> How should assessment be approached towards works in others genres or having different performance conventions such as jazz, minimalism, world music... how do we address performances of pieces that are considered 'too easy' or 'too hard'... how are performances involving unusual or unfamiliar instruments to be assessed...? (p.31).

The scheme that is used for the activities in the *Baroque Studies* and *Renaissance Studies* modules differs from that used for *Performance Studies*. The central activity

consists of seminars that involve group presentations. These groups normally comprise four students, determined by the tutor. Assessment is broadly two-fold: intra-peer and inter-peer.

Firstly, a confidential self and group assessment report is provided by individual members of the group in which they conduct an intra-peer assessment of their colleagues. The assessment criterion for this is holistic in character, being based upon the identification of individuals' contribution to the preparation of the seminar topic. Each group member submits their marks and these are totalled and an individual mark is allocated to each student following the 'zero-sum' method, as explained earlier in this chapter.

Secondly, each group's seminar presentation is inter-peer assessed by the other 'listening groups' each submitting a report and mark. The average mark is obtained from these, which is then compared with the tutor's assessment. The final mark is simply the average of the tutor's mark with that of the students. The tutors reported that marks awarded by the peer groups were found to be generally trustworthy.

As in the *Performance Studies* modules, the tutor, on the basis of representing a range of abilities and a mix of personalities, similarly determined the groups. Initially, students preferred tutor-determined groups, although there was scope for students themselves to decide on their own group formations, subject to tutor approval. The key issues involving the formation of groups, were considered in Sections 1.3.3 and 1.3.4. Although group work involving knowledge-centred music history modules may not require groups to be formed on the basis of instrumental expertise, using a mix of personalities as a source for determining the groups, however, emphasises interpersonal qualities. In such settings, students may learn much about themselves and about their group members in terms of their personal attributes and how these are displayed.

Assessment criteria for this module are discussed and negotiated by the students and tutor working together. The students typically generate non-weighted category criteria. Additionally, *Renaissance Studies* students (year two) are permitted to negotiate the focus, marks and weighting of their written assignments. As such, it would appear

considerable latitude is given to students in encouraging and allowing them to decide on the assessment process.

Another music history module *Trends in Late 19th and Early 20th Century Music* adopts a peer assessment scheme. As before, students were allowed to develop the assessment criteria, which were devised along typical non-weighted categories. The broad areas of criteria that students identified included structure, content and presentation but also included issues such as responses to questions. Their experience, similar to that of Falchikov (1995a), was that students' comments tended to focus on presentation rather than on content. Was this partly as a consequence of the assessment criteria that were used however? The 'responses to questions' criterion would have encouraged students to demonstrate a facility for argument, understanding and a capacity for deeper, rather than surface, learning, however. The tutor suggests that although 'assessing argument and understanding is particularly challenging for students' he finds by 'improving training, refining topics and introducing a requirement to ask questions...encourage[s] students to make comments that move beyond the presentational' (p.72).

It was reported that students tended to mark each other a little too high, marks falling within a narrow range, and did not 'enjoy the idea of marking down a peer' (Falchikov, 1995a, p.177). A few students were concerned about assessments that were biased, but this seemed to have reflected social groupings within the cohort. Following on from the previous discussion about group formations in *Baroque and Renaissance Studies*, would a tutor also need to consider the likelihood of biased marks occurring as a consequence of the group mix? They also found that students' unfamiliarity with the scoring system led to a lack of consistency in applying marking standards. However, assessments conducted by groups or panels were more objective and accurate than those produced by individuals, and the use of group-based marks was the preferred assessment method.

Overall, the major issue that emerges from the use of peer learning and assessment in this module is concerned with students' perceptions of the body of knowledge that seemed to be required in the course. Acquiring this knowledge encouraged several students to focus on the more surface, superficial and factual elements - both in their seminar presentations as well as the peer assessment of these. This seemed to be at the

expense of acquiring a deeper approach to knowledge through demonstrating, for example, their understanding of ideas and questioning of concepts and arguments.

The final activity described by Ulster University in the Peer Learning In Music project is a case study in the teaching of orchestration. It focused on the second and third years of a three-year project evaluating the usefulness of peer assessment and feedback in modules involving orchestration activities. The training session, which lasted one hour, included an explanation of the process and the development of assessment criteria by students in discussion with staff. As on other previous occasions, tutors, on the basis of a mix of abilities and instrumental expertise, selected the inter-peer assessment groups. It was again found that group assessment and group feedback proved more helpful than individual assessments. The groups completed a report on each orchestration assignment awarding each a percentage mark, which were subjected to tutor moderation. In general it was found that the student marks 'were remarkably close to those agreed by the tutor' and all feedback reports 'identified some key strengths and weaknesses' (Hunter & Russ, 2000, p.94). It was reported that some assessing groups were a little less assured and rather vague where sophisticated judgements were called for. There were also doubts from several students including concerns about their expertise, uncertainty about what they were doing, judgemental inexperience, and discomfort with the responsibility of marking. Some commented on the length of time that was taken with assessment and there was 'almost universal agreement that the process was hard' (Hunter & Russ, 2000, p.96). However, in contrast with *Trends in Late 19th and Early 20th-Century Music* there was an absence of biased marking in the orchestration modules. The study identified a need for more guidance on working in groups and distribution of work.

The work at Ulster University significantly advances our knowledge, not only of peer assessment in music, but has applications that can inform peer assessment across subject disciplines generally. It is a major contribution to the literature of peer assessment.

1.5.3 James Cook University

Daniel (2004) outlines a peer assessment procedure within a music performance context involving inter-peer assessment of solo performances. Students at James Cook

University were asked to assess as individuals, rather than grouped together in panels. Following the development of assessment criteria, students were also required by the tutor to identify weaknesses and strengths as a part of the written assessment. The weaknesses were later changed to 'areas requiring the most attention' (p.96). Before being returned to the performers, the written feedback comments were checked by the tutor. Some of Daniel's paper is concerned with quantitative assessment comparisons but he also conducted a detailed questionnaire in order to obtain students' responses to peer assessment. Although some students were not motivated by peer assessment, he concluded that peer assessment can lead to 'the provision of a more holistic performance feedback environment for students' (p.107). He also concluded that there was a need to consider students' assessment over time in order to examine the extent in which peer assessment contributes to the overall development of students' critical assessment skills. His findings suggest that students are initially less critical than tutors, in terms of marks, but more so later on, as they develop assessment knowledge and become increasingly confident about peer assessing. One area that might have been explored further, developing strengths and weaknesses, was not pursued and this aspect attracted surprisingly little comment in the paper.

1.5.4 University of Western Sydney

Blom and Poole (2004) discuss a project at University of Western Sydney involving inter-peer assessment of solo performances in which third year undergraduates were asked to assess their second year peers. The writers briefly explained the idea of using the third year students acting as assessors because they were already engaging with forms of assessment including self-assessment and peer discussion of music. Their preparation or training for peer assessment, however, appeared to be minimal, consisting of a single lecture 'discussing different approaches to the assessment process and reminding the students of how we, the performance staff, undertook our assessment of them' (p.111). Students were also asked to use criteria for assessment that were 'adopted by the staff assessment panel' and in so doing omitted the learning opportunities that might have been available through involving students in discussing and generating criteria for themselves. The criteria were based on: confidence/stage presence; musical communication; repertoire; awareness of style; playing technique. Students provided written comments, however, which proved interesting. They reported

a number of problems, a major difficulty being that of criticising peers constructively. Also, they found it problematic when assessing performances on instruments in which they had performing knowledge and there was a tendency to make a harsher assessment of these instrumentalists. The difficulty of assessing without a score was noted. The student assessors also addressed comments about their peers and were asked to give marks out of 15 for each performer. Although the marks did not count towards a summative assessment, they were not discussed with either the staff or the performers. If so, was an opportunity for learning from this marking feedback lost? The written comments about the performers were addressed to the staff; this raises the question posed earlier concerning written feedback about students being accessed by authority figures. The authors concluded, however, that the students learnt to critically evaluate the performing of others across areas of assessment and can be prepared for the roles of assessor and critic. The findings also suggest students have the capacity to be trusted and to evaluate their own learning. Indeed, their written responses indicate that many, to borrow a term used by Ramsden (1992), were engaging in 'deep' rather than 'surface' areas of learning.

Summary

This section has surveyed and discussed key studies in the literature on peer assessment in HE music curriculum. Among the issues reported by Kingston University were that lower levels of approval were caused by an unclear rationale for peer assessment. One benefit, however, was in the development of students' critical analysis. The Peer Learning in Music Project coordinated by the University of Ulster emphasise the centrality of learning and peer training in many of its various contributions. Although performance based group work and popular music are areas that are not specifically addressed, the Project is an impressive resource for practitioners and adds significantly to our knowledge of peer assessment in music. Daniel's paper involves quantitative marking comparisons at James Cook University, which suggests that students' marking abilities improve as they develop assessment knowledge. Although some students were not motivated by peer assessment, it did lead to improvements in feedback. The study at University of Western Sydney suggested that students had the capacity to be trusted and to evaluate their own learning. However, it appeared that several learning opportunities for students to learn from peer assessment feedback were lost.

Overall, the predominant musical activity in the literature was that of solo performance. Group work was uncommon, as was peer assessment in popular music idioms; these key areas are investigated in this thesis.

1.6 Epilogue: The Claims and the Pains of Peer Assessment

This chapter has surveyed the key studies that were thought relevant to the enquiry and subject area of this thesis. The survey has attempted to address those areas in the literature on peer assessment considered to be of particular significance: learning; group work; process features; the music curriculum. One interesting conclusion that may be drawn from this survey is that surprisingly few papers examined peer assessment activities in terms of its appropriateness in assessing particular learning activities or contexts. The nature of the curriculum activity and its potential for using peer assessment, in order to support particular aspects of learning, is rarely made explicit in the literature. This issue is a central concern of the enquiry and will be revisited in each of the remaining chapters.

This final section is an epilogue to the chapter, rather than being a further exposition. As most of the studies within the literature on peer assessment report a variety of benefits or problems that have been encountered, I thought it useful to assemble a list of these, which I have entitled 'claims' and 'pains'.

Falchikov (2005) as a part of her conference presentation about peer assessment in HE, for example, provided a list of the benefits of peer assessment that were organised into a number of categories. As much of her writing is enthusiastic and supportive of peer assessment, it is perhaps unsurprising that she did not present an equivalent assemblage of its problems. The categories that were used by Falchikov in order to present the benefits have been adopted in Table 1.1 (the 'claims') and I have updated and extended the literature sources that she originally provided.

A corresponding 'pains' of peer assessment is illustrated in Table 1.2. The categories here were developed from the literature that was surveyed.

Table 1.1: The Claims of Peer Assessment (adapted from Falchikov, 2005)

Claims of peer assessment	Examples
Cognitive and meta-cognitive competencies	
Aids problem solving	Dochy *et al.* (1999) (review)
Encourages reflection*	Boud & Knight (1994) Falchikov (1996; 1998) MacDonald (2000)
Encourages transfer of learning	Catterall (1995)
Improves critical thinking*	Falchikov (1986) Oliver & Omari (1999) Sivan (2000) Smith *et al.* (2002)
Improves understanding/mastery*	Catterall (1995) Falchikov (1986) Lapham & Webster (1999) Ney (1989) Searby & Ewers (1997)
Brings about unspecified educational/learning benefits	Davies (2003a) Sitthiworachart & Joy (2004) Keppell *et al* (2006)
Skills development	
Enhances listening skills	Falchikov (1995a & b)
Enhances vocational skills*	McDowell (1995) Trevitt & Pettigrew (1995)
Improves presentation skills	Price & Cutler (1995) Brindley & Scoffield (1998)
Improves writing skills	Topping (1998) (review)
Promotes learning skills/abilities*	Dochy *et al.* (1999) (review) Ballantyne *et al.* (2002)
Promotes evaluative skills*	Searby & Ewers (1997) Ballantyne *et al.* (2002) Bryan (2006) Smith *et al.* (2007) Fallows *et al.* (2001)
Promotes lifelong learning skills*	Challis (1999) Ballantyne *et al.* (2002) Boud & Falchikov (2006)
Brings about unspecified benefits to professional skills*	Topping (1998) (review) den Berg *et al.* (2006) Pond & ul-Haq (1997) Heylings & Stefani (1997) Strachan & Wilcox (1996)
Performance	
Enhances experience of trainee teaching	Anderson & Frieberg (1995)
Improves academic performance	Dochy *et al.* (1999) (review) Davies (2003b) Bryan (2006) Mowl & Pain (1995) Tsai *et al* (2002) Bhalerao & Ward (2001)
Improves grades/test scores	Bangert (1995) Dochy *et al.* (1999) (review) Shortt (2001)

Table 1.1: The Claims of Peer Assessment (continued)

Personal/intellectual development	
Improved awareness of quality of own work*	Falchikov (1995a & b) Freeman (1995) Mehrens *et al.* (1998) Smith *et al.* (2007)
Increases autonomy	Falchikov (1986) Lapham & Webster (1999) Sivan (2000) Stefani (1998)
Increases responsibility*	Dochy *et al.* (1999) (review) Edwards & Sutton (1991) Lapham & Webster (1999) Heathfield (1999)
Increases self-efficacy (in the context of mathematics)	Bangert (1995)
Brings about unspecified benefits	McDowell (1995)
Social competencies	
Enhances diplomatic and negotiation skills*	Falchikov (1994; 1995a & b) Topping *et al.* (2000)
Improves cooperation*	Lapham & Webster (1999) McDowell (1995) Orsmond & Merry (1996) Smith *et al.* (2007)
'Affective dispositions' (Birenbaum, 1996)	
Decreases test anxiety (particularly mathematics anxiety)	Bangert (1995)
Increases confidence*	Lapham & Webster (1999) Price & Cutler (1995) Sivan (2000)
Improves internal (intrinsic) motivation*	MacDowell (1995) Oliver & Omari (1999)
Reduces stress	Zakrzewski & Bull (1998)
Benefits to assessment	
Enhanced appreciation of importance of criteria*	Trevitt & Pettigrove (1995) Ballantyne *et al.* (2002) Bloxham & West (2004)
Develops reliable standards of peer marking	Billington (1997) Rudy *et al.* (2001) Oldfield & McAlpine (1995) Orpen (1982)
Brings unspecified benefits	Davies (2002) Hanrahan & Isaac (2001)
Benefits to tutors	
Saves time	Miller & Ng (1996) Hanrahan & Isaacs (2001)
Assesses course outcomes	McGourty *et al.* (1998)
Unspecified (associated with the need to prepare model answers	Gray (1987)

Table 1.2: The Pains of Peer Assessment (categories emerging from the literature)

Pains of peer assessment	Examples
Unclear or uncertain tutor involvement	
Tutor role unclear to students	den Berg *et al.* (2006)
Not all tutors committed	Ballantyne *et al.* (2002)
Students regard assessment exclusive responsibility of tutor	Brindley & Schofield (1998) Davies (2000)
Student preference for tutor to moderate marks	Ballantyne *et al.* (2002)
Student preference for marking by experts	Searby & Ewers (1997)
Doubt about awarding marks	
Students feel unqualified to mark each others' work	Orsmond & Merry (1996) Sullivan *et al.* (1999) Li (2001) Sluijmans *et al.* (2001)
Unclear about criteria for awarding marks	Ballantyne *et al.* (2002)
Doubts concerning accuracy of peer marking	
Reluctance to award low marks	Brindley & Schoffield (1998)
Bunching/narrow range of peer marks	Magin (2001) Hughes & Large (1993) Stefani (1994) Freeman (1995) Sullivan *et al.* (1999) Purchase (2000) Miller (2003)
Poor correlation of students/tutors marking	Swanson *et al.* (1991) Freeman (1995) Mowl & Pain (1995)
Naïve and inaccurate marking*	Swanson *et al.* (1991) Orsmond *et al.* (1996) Miller (2003) Rust *et al.*(2003)
Student inexperience or credibility	
Sceptical about the worth of peers' comments	Orsmond *et al.* (1996)
Not supportive of first year students being involved	Cheng & Warren (1997)
Concerns about ability to provide constructive feedback*	McDowell (1995)
Time consuming	
Time consuming process for students	Orsmond *et al.* (1996) Davies (2000) Topping *et al.* (2000)
Labour intensive for tutors	Cheng & Warren (1997) Searby & Ewers (1997) Race (1998) Brindley & Schofield (1998) Ballantyne *et al.* (2002)

Table 1.2: The Pains of Peer Assessment (continued).

Invites questionable or unethical practices	
Bias and collusion*	Brown & Knight (1994) Pond *et al.* (1995) Ballantyne *et al.* (2002) Sharp (2006)
Friendship marking*	Pond *et al.* (1995) Cheng & Warren (1997)
Student fear of their grade adversely affected by free-riders*	Abson (1994) Nicholson & Ellis (2000) Doran *et al.* (2000) Keppell *et al.* (2006)
Deliberate attempts to manipulate marking system	Parsons & Drew (1996) Lejk & Wyvill (2001b) Bushell (2006)
No confidence in fair marking	McDowell (1995)
Marking on personality/popularity rather than on intellect	Doran *et al.* (2000)
Gender bias	Langan *et al.* (2005)
Damaging for group cohesion	
Dislike of criticising friends*	Williams (1992) Straughan & Wilcox (1996) Hanrahan & Isaacs (2001)
Adverse affective dispositions	
Causes stress	Pope (2001) Pope (2005)
Disputed assessment procedures	
Students do not take it seriously*	Ballantyne *et al.* (2002)
Students not taking it seriously for formative assessment	Swanson *et al.* (1991)
Concerns about legal challenges	Doran *et al.* (2000)
Complaints about assessment without preparatory training	Cheng & Warren (1997) Sluijmans *et al.* (2001)
Unwieldy methods of identifying individual contributions using peer assessment	Conway *et al.* (1993)
Reliability and validity concerns	O'Donnell & Topping (1998)
Concerns about peer assessed grades contributing to degree classifications*	Doran *et al.* (2000)
Poor student contributions benefiting from group marks	Pond *et al.* (1995)

The purpose of assembling these tables, however, is one that goes beyond a meta-analysis of the claims and pains in the literature: identified here, are a number of categories (indicated by an asterisk) thought to be of particular relevance to this enquiry and that represent many of its key concerns. For the 'claims' of peer assessment, these would include the following general categories: *cognitive and meta-cognitive competencies; skills development; personal/intellectual development; social competencies; affective dispositions.* For the 'pains' of peer assessment, the overwhelming concerns were thought to belong to the following two general categories:

invites questionable or unethical assessment practices; disputed assessment procedures. The implications for this research, arising from these 'claims' and 'pains' that gradually became manifest as the enquiry proceeded, are, examined, analysed and clarified, especially in Chapters Five, Six and Seven.

Chapter Two: Methodological Considerations

2.1 History and context

2.2 Considering paradigmatic assumptions

2.3 Considering peer assessment: exploration, practice, issue and questions

2.4 Rehearsing and performing cycles

2.5 Considering mixed methods

2.6 Considering an action research approach

2.7 Role and ethnographic considerations

2.8 Consideration of values

2.9 Supporting learning

2.10 The curriculum area

2.11 Summary

2: Methodological Considerations

It is the intention of this chapter to offer an account of the various methodological considerations that emerged during the shaping of my research ideas. A purpose of these *Methodological Considerations* is to provide a contextual background to Chapter Three: *Methods of Enquiry*. Where this relationship between both chapters is particularly explicit, this is indicated by a reference to the equivalent section in *Methods of Enquiry*. Similarly, *Methods of Enquiry* contains a number of references, where appropriate, back to this chapter.

I have organised *Methodological Considerations* into sub-sections in order to identify, and make clear, the key considerations that helped to inform and shape the development of the methodology of the research enquiry.

2.1 History and context
2.2 Considering paradigmatic assumptions
2.3 Considering peer assessment: exploration, practice, issue and questions
2.4 Rehearsing and performing cycles
2.5 Considering mixed methods
2.6 Considering an action research approach
2.7 Role and ethnographic considerations
2.8 Consideration of values
2.9 Supporting learning
2.10 The curriculum area
2.11 Summary

The interconnectivity between these areas became apparent as the ideas for the research emerged; Figure 2.1 illustrates this (p.71). These are the key 'methodological considerations' of this chapter; together they offer a flavour of the kind of personal journey that was undertaken in the development of the ideas contained in this thesis.

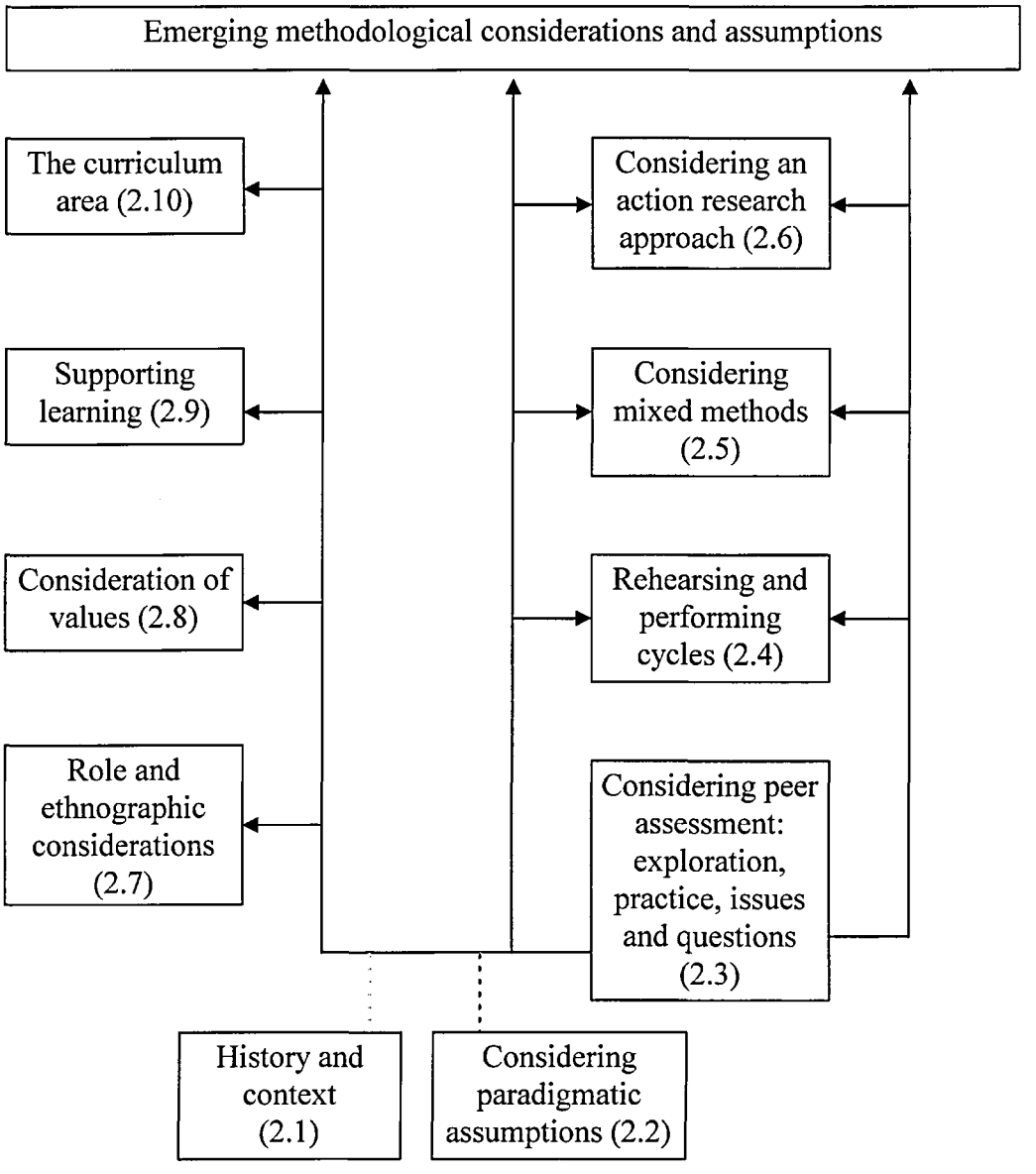


Figure 2.1 Emerging methodological considerations and assumptions and their interrelationships

The first section *History and context* (2.1), although *not strictly concerned with considerations of a* methodological *character*, explains the origination of the enquiry and considers its situational setting, together with an outline of the context in which popular music undergraduate students were involved in the research.

Questions of paradigm and assumptions made about the nature of knowledge and their influence on this research are explored in *Considering paradigmatic assumptions* (2.2). This section discusses the emergence of the ontological and epistemological positions, which, I considered, informed and underpinned the methodological design. Along with

History and Context (2.1), its supportive nature is indicated by dotted lines.

Considering peer assessment: practice, issues and questions (2.3) describes the developing of my practice towards acquiring the necessary skills and understanding of the various operational mechanics and processes of peer assessment. A number of issues that are raised here may have a wider relevance for methods that, in particular, are intended to develop aspects of peer learning.

As I acquired confidence in using peer assessment techniques for performance-based modules, I began to explore its application to a wider context. This focused upon the activities of band rehearsals and performances. These successive activities – a number of band rehearsals followed by a performance - were often repeated as a part of the course. The nature of these activities or student experiences - rehearsals culminating in a performance - was of interest to me in the exploration of peer assessment. I decided to describe or entitle this sequence of activities as a *Rehearsing and performing cycle* (2.4).

As methodological considerations were evolving and emerging, I considered the use of multiple methods of data gathering would be appropriate. The section *Considering mixed methods* (2.5) examines reasons for adopting such an approach.

The type of enquiry that I was proposing appeared to suggest that an action research methodology was appropriate. The development of this is described in *Considering an action research approach* (2.6).

Underpinning these emergent methodological considerations were assumptions that are made about knowledge and how this is created through human interaction. As tutor and researcher, it was necessary to consider these roles and the impact of each. *Role and ethnographic considerations* (2.7) therefore examines these roles and the nature of ethnographic approaches that may be implicit in an enquiry of this sort.

Consideration of values (2.8) examines my values, in terms of learning, teaching and peer assessment, that I considered were relevant to the enquiry. Personal values have significance in action research investigations and these are considered in this section.

Supporting learning (2.9) in *The curriculum area* (2.10) of popular music is the *raison d'etre* for this enquiry.

These then are the key 'methodological considerations' of this chapter. Together they offer a flavour of the kind of personal journey that was undertaken in the development of the ideas contained in this thesis.

2.1 History and context

Although not strictly concerned with considerations of a *methodological* character, it may be useful at this point, to illuminate further the contextual background to the enquiry. The purpose of this section is that of providing a flavour of the courses, the student experience context, the subject area and the interface with peer assessment.
I think it is also helpful here to capture something of the character of the enquiry. In a few words, it can be described as an investigation into peer assessment approaches that assist learning in popular music group rehearsing and performance. Before continuing with this section, it may be helpful to be reminded of the research focus that developed during the enquiry: the contribution of peer-assessed personal attributes to student learning within the context of band rehearsing and performing activities

As briefly outlined in Chapter One, first year cohorts of students from the *Performance Management* module of the BA Popular Music Studies and the *Performance Studies* module of the BA Extended Degree course were involved in this enquiry. The starting points of the journey into peer assessment arose in the months prior to the formal commencement of the research. In the autumn of 1999, I decided to invite *Performance Management* students to peer assess their band members' contributions to group rehearsing, using the 'zero-sum' method (explained in Chapter One) following the experiences of Hunter & Russ (1996). I also thought that exploring peer assessment of band members' contributions to rehearsing might be an interesting research topic.

At the same time, I considered inviting third year *Performance Techniques* students (a module option within *BA Popular Music Studies*) to participate in peer panels in order to help assess the live gigs given by the BA1 Performance Management students.

Participation would be voluntary; being a panel member was not a required element of their course and no pressure would be applied. The voluntary nature of students' involvement in inter-peer assessment panels is discussed in Section 3.4.1.
Three factors had a bearing on deciding to involve year three performance students as inter-peer assessors.

(i) Firstly, bands would benefit from receiving assessment feedback from senior students with considerable experience in terms of performance ability and understanding of performance standards. I hoped that the involvement of senior performance students as panel members would provide weight to the peer assessment feedback, which might lead to this feedback being received more seriously.

(ii) Secondly, I assumed that the use of year three students to assess year one bands might help to minimise occurrences of friendship marking or other unethical practices, noted in Chapter One. Although friendship groupings or personal conflict might exist across cohorts, I assumed that these would be less probable between first and third year students, than within the same year.

(iii) Thirdly, senior students, having themselves successfully experienced the group performance modules, would more easily be able to understand this rehearsing and performance context and empathise with students' experiences.
Furthermore, they had already, themselves, received peer feedback as part of the weekly performance workshops, known as the 'rota', operating for year three students. The 'rota' consisted of individuals and bands performing to a peer audience who would provide verbal feedback. I hoped that this experience would also help to develop their evaluative and feedback skills, as training preparation for belonging to a peer panel. The 'rota' would be a key part of the peer assessment-training programme and this is described in Section 3.4.4.1.

A brief discussion about the course learning outcomes of each module concludes this section. It was important to establish that the objectives of the research would be supportive of the aims and learning outcomes of the Performance Management and Performance Studies modules. Particularly relevant for this enquiry was the

Performance Management learning outcome of the module that involved the ability to 'critically evaluate the group's management of rehearsals and performances and the performance of others.'[1] Similarly, I considered that this enquiry would be broadly sympathetic to the module aim that involved the 'development of performance skills, management abilities and collaborative group work.'

The aims and learning outcomes for the Performance Studies module were comparable with those of Performance Management. I thought this enquiry would be especially supportive of Performance Studies aims that involved 'performing the role of the musician in an ensemble, working effectively within a range of performing contexts' and developing the ability for 'critical evaluation of group rehearsals and performances.'

The third year Performance Techniques module was a popular option for advanced instrumental performance students. Students received a weekly individual instrumental or vocal lesson and all students met for a weekly performance workshop. As explained earlier in this chapter, the training that would support peer panels, in their assessments of band performances, was provided within these workshops. Overall, I concluded that my proposed research into peer assessment would support these aims and learning outcomes and, would assist individuals in achieving them.

As already explained, this part of the chapter has been concerned primarily with describing the background to my research interest in peer assessment, rather than with matters methodological.

2.2 Considering paradigmatic assumptions

This section addresses questions of paradigm and considers a number of assumptions about the nature of knowledge that informed the development of the methodology. Central to this were assumptions that were made about learning within group work. As indicated in Section 2.1, I considered that the context of the activity helps to define the kinds of learning that are intended. If learning requires a context, then it is essential to be able to understand and explain the particular learning context in which this study is

[1] This and subsequent extracts from module aims and learning outcomes of the modules appear in Appendix 1 (p.271).

placed. The purpose of this section then is to address questions and issues arising from assumptions about the nature of knowledge and the learning context in this enquiry. Students working together in their bands, rehearsing and preparing for public performances, typified the learning context for the modules. A fuller description of these activities can be found in Section 3.1.1. These activities, typically occurring in the modules, also reflected the 'real world' of the professional (described in Section 2.1). For this context, therefore, opportunities for learning were situated within natural settings: students' normal experience occurring in the modules and that also of the real world.

Understanding the nature of what was being learned in a band rehearsal situation was important for this enquiry. In order to address this from an epistemological position, I am making an assumption that knowledge can arise from the complexities of interaction among band members during their activities of rehearsing, performing and peer assessing. Furthermore, the context of this interaction is such that it can embrace both social learning experiences as well as the knowledge domain of popular music group rehearsing and performance.

This is similar to Glaserfield's (1991) view of knowledge as an activity or process rather than a particular kind of product that exists independent of the knower. I have previously suggested that the nature of acquiring such knowledge is one that might be described as being contextually embodied. Putting it another way, knowledge and the activity of learning is dependent upon a setting, a context. It is difficult, for me, to conceive knowledge and learning existing in some kind of remote cyberspace. Making sense out of social interactions, such as those occurring within a band rehearsal and peer assessment, together with apprehending the ontological realities of that encounter, surely depends upon the intellect and meta-cognition of individuals and their bands. Being the students' tutor, I considered myself, inevitably, as being included in such a contextual embodiment. This suggests that I was also a part of the social interaction in which learning and knowledge creation might occur. From this perspective therefore, I considered that my enquiry would embrace an ontological and epistemological position that involved my roles as tutor and researcher together with those of the students being interactively linked in the social and naturalistic context of the research, and that the

'findings are literally created as the investigation proceeds' (Guba & Lincoln, 1994, p.111).

This places my research, in terms of a basic belief system, within a social constructivist paradigm, following the account of Guba & Lincoln (1994) and Bruner (1986). Constructivism can be explained in part by Schwandt (1994), who believes that

> knowledge and truth are created, not discovered by mind... [and]...human beings do not find or discover knowledge so much as construct or make it. We invent concepts, models, and schemes to make sense of experience and, continually test and modify these constructions in the light of new experience (p.125).

To summarise, this section has explored paradigmatic assumptions about the nature of knowledge, its construction, and its context. Context is particularly important to this enquiry, being located within a natural Higher Education setting typified by normal activities and student learning experiences. These activities were characterised by social interaction: group rehearsing, performing and peer assessment, with myself, as the students' tutor, being interactively linked. Putting it simply, I am assuming that knowledge can be created through participation in peer assessment activities, and such knowledge arises from and belongs to those involved in its creation, importantly to the students, but in part also to myself, as their tutor.

At this stage, it is proposed to explore certain paradigmatic and epistemological assumptions related to the knowledge domain of popular music group performance. The following brief account examines a number of constructs that I considered were important, for this subject discipline.

Humans can make sense of their world through communication. Human communication has become manifest in innumerable ways including verbal and non-verbal, signs and symbols. Through these various verbal and non-verbal communication systems we create many languages in which we strive to understand and give meaning to our existence. One of these languages that we have constructed is that of organised sound: popular music, for example. The world of music, and other worlds, however, only properly becomes meaningful within a shared system of intelligibility (Gergen, 1991; 1995). Critically, for us to be able to also share these meanings and to construct or add to our knowledge of the language of music, we need to comprehend this world (Aspin,

1984). Understanding these complex meanings might necessitate, for example, being immersed in a HE music education experience, such as the *Performance Management* module. These assumptions, drawn from this brief account of meaning theory and philosophy of language are discussed more fully, in relation to the completed action research, within Chapter Seven. They are raised here, however, because such constructs have relevance to the social learning contexts involved in popular music group work.

In adopting a social constructivist view and being persuaded by the arguments articulated by Goodman (1978), Bruner (1986), Goodman & Elgin (1988) and Fosnet (1996), the participants of this research would be engaging in a social context of group music making rooted in real world contexts; not somehow engaging with an external world or artificial experiment. Indeed I regarded that the creation of popular music group work knowledge among bands would begin with 'what happens to be currently adopted and proceed to integrate, organise, weed out and supplement' (Goodman & Elgin, 1988, p.163).

Although several assumptions that involve this research enquiry have already been examined, there are several other issues remaining. These involve considering certain paradigmatic assumptions that are made about the following: my role as researcher and tutor, interviews with students; and verification and quality of research findings.

2.2.1 Paradigmatic assumptions: research role and tutor role

Being both tutor and researcher raises questions about my role in this enquiry. I believed that this investigation would necessitate a prolonged engagement with students, as a tutor and as a researcher, extending to five years. Furthermore, I believed that it would be difficult to be able to disentangle myself from each role in order to be completely detached from rehearsing activities and student experiences. I considered it inevitable that, by being involved with students as their tutor and researcher, I would contribute to the student experience.

Being a part of these activities resonates with constructivist theory: Guba & Lincoln (1994), for example, in their account of constructivism, not only believe that the

observer cannot be separated from the observed in the activity of enquiring into knowledge constructions, but also that they should not be separated from it.

2.2.2 Paradigmatic assumptions: the interview

I considered at an early stage in the planning that conducting interviews with students would be an appropriate data gathering method with which to obtain their responses to peer assessment. Silverman (1993) explains two kinds of approaches, positivist and interactionist, in which there are underlying paradigmatic assumptions for the research interview.

A positivist approach assumes that the primary purposes of interviews is for obtaining factual information, which may be tested for reliability and validity and that those facts exist independently from the research setting. Pre-determined research protocols and methods such as structured interviews and standardised questions are typical of interviews influenced by a positivist research paradigm.

An interactionist approach, however, assumes that the interview and interview data is constructed by the encounter between the interviewer and interviewee. The interview, in other words, is context-laden and the interaction between its participants contributes both to the data and to its interpretation. Indeed, being inseparable from the observed activity, through my participation in the interviews, I considered was congruent with constructivist thought.

The emphasis of the interactionist interview lies in exploring and understanding the experiences, behaviours and attitudes of the interviewees, rather than in obtaining facts, pure and simple. Pre-determined research methods may not always be appropriate for an interactionist approach, where interviews are typically un-structured, semi-structured or open-ended. I concluded that an interactionist approach in which a combination of open-ended and more focused questions would be appropriate for this enquiry. Sections 3.5.1, 3.5.2 and 3.5.3 explain how I developed the interview format that I considered was suitable for the investigation.

2.2.3 Paradigmatic assumptions: verification and quality of research findings

In placing my investigation within a constructivist paradigm, I considered that the need for the verification of my research and for evaluating its quality might be problematic in terms of establishing appropriate criteria. Whereas in positive thought, for example, the emphasis is on the conventions of internal and external validity, reliability and objectivity, constructivists do not believe that there are unquestioned foundations for any interpretation. Lincoln & Guba (1985) propose criteria that encompass *trustworthiness.* Trustworthiness includes credibility (paralleling internal validity), transferability (paralleling external validity), dependability (paralleling reliability) and confirmability (paralleling objectivity). Their description of *trustworthiness* is an attempt to provide criteria for judging the quality of constructivist research and has been well received. For my research, I considered that adopting a *trustworthiness* concept of quality, following the account of Lincoln & Guba, was appropriate and consistent within the constructivist paradigm that the enquiry would embrace. A description of the criteria for the standards of verification and quality of this research appears in Chapter Three.

This section, *Considering paradigmatic assumptions,* offers an exposition of the many assumptions that were made in shaping my research ideas and their development. I now turn to considerations of more pragmatic matters that involve a journey of learning the nuts and bolts of peer assessment.

2.3 Considering peer assessment: exploration, practice, issues and questions

My encounters within the world of peer assessment during the first two years of this research could be described as an exploration. Important in this, however, was my belief that for the research to be founded upon a basis of good practice, it was necessary for me to acquire extensive peer assessing experience. The purpose of these first two years therefore, was twofold. Firstly, it enabled my own practice to develop through acquiring skills and also confidence in being able to operate effectively both as a tutor and as a researcher. Secondly, it established the foundations and the direction of the peer assessment enquiry. Arising from this were three areas of enquiry: inter-peer assessment, intra-peer assessment and the process of peer assessment. Within each of these I identified a number of key issues, which I felt were of particular interest. Some

of these issues did not appear to be raised or explored in the published literature on peer assessment. An overview of the issues and questions that I considered of interest appears in the following sub-sections.

2.3.1 Inter-peer assessment

I began first by examining the assessments given by the peer panels (described in Section 2.1). This involved comparing the marks awarded by the peer panels with those of the tutors and exploring the process of inter-peer assessment. I also gathered qualitative data arising from three interviews with the peer panel students. Two questions began to emerge which are explored further in Chapter Three.

- What do peer panel students themselves learn during their inter-peer assessing experiences?
- How accurate and consistent are peer panel marking judgements compared with those of the tutors?

2.3.2 Intra-peer assessment

A second area involved examining ways in which the students' individual contributions to the group rehearsing might be identified. I explored a number of approaches to intra-peer assessment of group rehearsing, in order to evaluate an individual's contribution to group rehearsals. As for inter-peer assessment, I found the marking aspect, particularly in terms of the consistency of students' agreements, similarly fascinating. Arising from this was the identification also of free riders and shooting stars within the bands. However, I became increasingly attracted to exploring how peer assessment might be used to encourage students to learn more about themselves and about each other during their rehearsals. Of particular interest were the issues raised by Pearce (2000) and others (discussed previously) that the development of personal attributes ought to be a central concern of music education.

I considered how personal attributes of the popular music students might develop through using an intra-peer assessment approach. Within group rehearsals, it seemed to me that students' personal attributes, particularly those relating to their teamwork

competencies, would be particularly important for productive rehearsals. I thought that two kinds of personal attributes might be developed: personal attributes relating to the group or band as a whole and specific attributes relating to each individual within the group. I considered that there were at least four questions that required to be addressed here.

- How might students' personal attributes in a band-rehearsing context be identified?
- Were there particular personal attributes that could be considered as being more amenable in terms of improvement?
- How might learners' personal attributes be developed or improved through the use of peer assessment?
- What can we understand from the feedback and marks that students gave?

Using students' personal attributes as intra-peer assessment criteria and the processes through which to provide opportunities for learning, discussed below, became the key focus of the research.

2.3.3 Peer assessment processes

A third area of the enquiry related to developing my practice, understanding and confidence in using this mode of assessment. I began to understand the various operational mechanics and the decisions, within the process, that were required. Emerging from this were the following three questions.

- Does the timing or sequence of activities within a peer assessment mechanism affect student learning?
- How does the setting of the intra-peer assessment, such as marking in private or in collaboration, affect learning?
- Does a peer assessment *system* (that incorporates formative, summative, intra- and inter-peer assessment activities, for example) have an influence on learning?

An outcome of the research arising from using students' personal attributes and the processes involved towards developing learning would be a new process model for intra-peer assessment; this is discussed in Chapter Six.

The data that was collected during the first two years helped to inform my understanding of the process features of peer assessment. Quantitative data arose from the assessments themselves: this data played a useful informational role within the initial exploration, methodological approaches and, ultimately, the implicit paradigm. Three interviews with peer panel members were also conducted in order to obtain their responses to peer assessment.

To summarise, *Considering peer assessment: exploration, practice, issues and questions* identified three key areas arising from this exploratory phase of the research. These were inter-peer assessment; intra-peer assessment; and peer assessment processes. A number of issues and questions emerged which would become the basis of further investigation. A prime purpose of the research during this period was of obtaining experience and understanding of peer assessment, through which I would be better equipped to progress the issues and questions of the enquiry.

These issues and questions within the three areas would continue to be explored throughout the enquiry. A chronological description and explanation of the investigations conducted into these three areas can be found in Chapter Four.

2.4 Rehearsing and performing cycles

As previously explained, the investigations were conducted in the students' natural setting (a Higher Education institution, during normal scheduled sessions). The central activity of the BA1 *Performance Management* module was that of making music in their band ensembles: rehearsing and then performing on a stage. There were four medium sized band rooms and bands were usually assigned one and a half hours of rehearsal time per week. They were supervised by myself and sometimes by another tutor at these times. Bands could also book additional rehearsing time that was unsupervised. As a tutor, I worked with each band in turn from room to room on developing their song material and providing feedback on their performing. This activity was repeated at various times during the module and, from year to year, with each successive cohort of students.

As such, I considered that this repeated pattern of rehearsing and performing could be described as a *cycle* of activities. I decided to use the term *rehearsing and performing cycle* to describe the typical pattern of these activities, in which peer assessment was a part. The cycle usually comprised the following activities: introduction and explanation of the assignment brief; establishment of assessment criteria; preparation or training activities; deciding band membership; rehearsing; intra-peer assessment; performance; inter-peer assessment; and feedback.

Each successive rehearsing and performing cycle added to my experience, improved my understanding and increased my confidence in peer assessment. The conclusion of each cycle was also a place at which I was able to reflect upon my personal practice as tutor and researcher[2]. It allowed, therefore, particular aspects of peer assessment to be explored, changed and compared from one cycle to the next. In this sense, the cycle of rehearsing and performing during the course module resembled that of an *action research cycle* as described by McNiff (1988) and McNiff *et al.* (1996).

Deciding on 'intervention and change' (Robson, 2002, p.219) in the rehearsing and performing cycles was guided by the literature and given focus by a set of values. Regarding the latter, I considered it important that my understanding of this mode of assessment and its potential for assisting learning was also congruent with my value position. This aspect is discussed more fully in Section 2.8.

2.5 Considering mixed methods

Methodological considerations during 2000-2002 were being gradually refined with each rehearsing and performing cycle. I believed that adopting a responsive and flexible action research design would be more appropriate to the enquiry than, for example, employing a formal, pre-defined methodological apparatus. The over-arching aim of this thesis involved exploring a number of different sub-questions and issues, rather than focusing upon a single research question. As such, I considered a rigid design structure as inappropriate due to it restricting unforeseen areas of enquiry that might emerge.

[2] An account of my reflection on personal practice can be found in Chapter Seven.

At an early point in the investigation I thought it likely that the different stages of the enquiry would require a flexible approach to data collection. Questions that were raised in Section 2.3 appeared to suggest the use of both quantitative and qualitative methods at certain times. For example, where the focus of the enquiry involved comparing peer panel marking with that of tutors, a quantitative approach would be required using peer assessment marking data. At other times, however, a qualitative approach might be more appropriate. Obtaining individuals responses to peer assessment would, for example, suggest gathering qualitative interview data.

This suggested a methodology that allowed for a combined approach using both qualitative and quantitative data. Using a mixed method of data collection would enable many aspects to be covered. I also envisaged that the chronology of gathering the various kinds of data using mixed methods would be overlapping during the investigation. Data overlap, for example, would occur if an interview with a Performance Management student had to be conducted during his or her second year, as a consequence of the practicalities in arranging it. Meanwhile, fresh data collection would have already commenced with the new cohort of *Performance Management* students. A methodological design that would be responsive to collecting and analysing mixed sets of data at particular times, required careful thought. The following section considers this aspect of the study.

2.5.1 Deciding on data organisation and reporting

Deciding on an appropriate structure for planning, reporting and presentation of overlapping quantitative and qualitative data, was problematic. The literature did not appear to offer any similar examples of practice in the area. However, two possible design structures were considered: segregation between quantitative and qualitative data reporting; or, reporting both data types together in a chronological sequence.

The first method would present separate qualitative and quantitative data reports. This would provide clarity for each separate data collection method. However, this design structure would not easily report the developmental and overlapping character of the data nor adequately account for relationships between the two data types.

The second method would present a chronological structure, involving data arising from the sequence of rehearsing and performing cycles. Sub-headings within the chronology would report the particulars of the data (for example, rehearsing and performing cycles 2003-2004; individual interviews with Performance Management students). A disadvantage of this design, however, is that it might lead to a fragmented presentation, requiring many sub-headings in order to report how the data arose from each successive rehearsing and performing cycle.

It was also important that a design structure for organising, reporting and presenting overlapping quantitative and qualitative data was appropriate for a naturalistic research of this sort. As explained already, this enquiry assumed that knowledge is constructed through the interactions and experiences of the students in a natural rehearsing and performing setting.

This setting indicated that a structure for reporting the data should follow the developing experiences of the students as closely as possible. I considered, therefore, that a chronological organisation would best report the developing experiences and responses of the students emerging in the data. Importantly, a chronological reporting and presentation would also indicate, for example, the interventions that would be made in the rehearsing and performing cycles. Chapter Four illustrates the structure that was used to report and present the data and interventions in the rehearsing and performing cycles.

Section 2.5 has discussed a number of key considerations in the development of the methodology: use of qualitative and quantitative methods, overlapping data collection, emerging designs, action research, a re-visit of constructivist theory and my researcher-tutor role. This section has also commenced a synthesis of some elements within the enquiry, which I considered were crucial, in ensuring that the integrity and central concerns of the investigation would be clear. The synthesis continues in the next section, which considers the particular action research approach that was developed.

2.6 Considering an action research approach

As previously explained, my explorations into peer assessment during 2000-2002 typified an action research approach. For example, I considered that the rehearsing and performing activities were characteristic of 'action research cycles' (McNiff, 1998). Also, McNiff might describe my participation, as researcher and tutor in the enquiry, as being 'insider research'; researching, analysing and reflecting on my own educational practice are activities that typify action research methodologies.

Related to the practitioner research aspect of this enquiry was the necessity for developing my understanding of peer assessment practice. I considered that acquiring experience in this was crucial for the progression of the research. It was also necessary to defend the accusation that, in my research role, I lacked the expertise and competence that was required (Carr & Kemmis, 1986; Guba & Lincoln, 1994).

Within the discussion of Section 2.5, *Mixed Methods,* a synthesis of various elements of the enquiry was commenced. Developing an action research methodology that was appropriate for the enquiry involved a further synthesis of its elements; this is illustrated in Figure 2.2 (p.88).

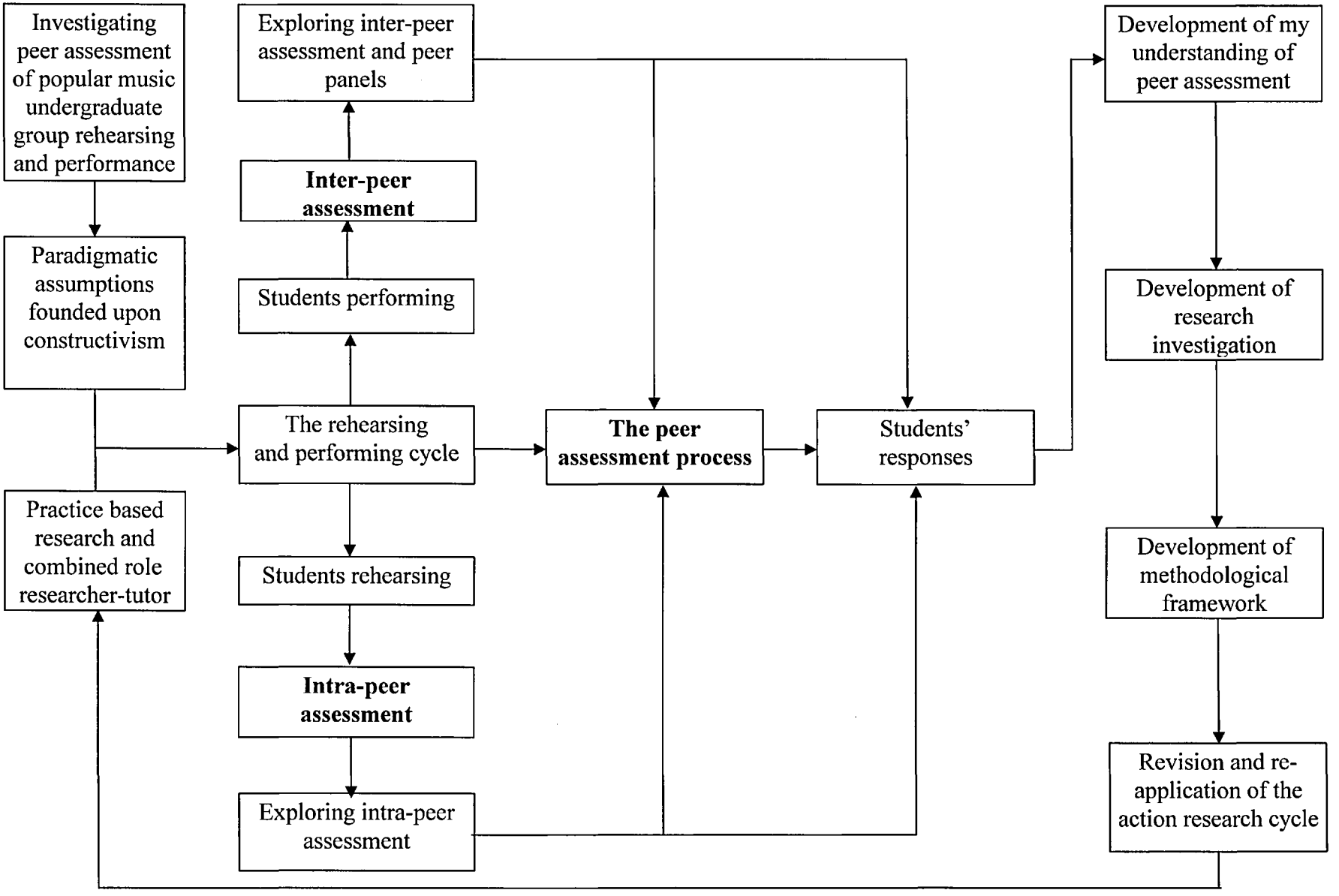


Figure 2.2 Synthesis of the elements of the enquiry

The key elements that I thought were central to the investigation and for the developing methodology are to be found in Figure 2.2. This diagram also illustrates the relationship between these elements.

In the introduction to this section, I explained that my exploratory investigations were characteristic of an action research approach. These characteristics may be noticed in the diagram: rehearsing and performing cycles of action; revisions to each cycle and re-application; and development of the tutor's knowledge and peer assessment practice. Figure 2.2 also indicates the three areas, discussed in Section 2.3, which became of interest to me (emboldened): intra-peer assessment; inter-peer assessment; and the process of peer assessment.

At the centre of the developing action research methodology were the subjects of the research: the students and their bands. As explained in Section 2.1, the subjects were undergraduates belonging to the year one Performance Management module, the foundation year Performance Studies module, together with peer panels drawn from the year three Performance Techniques module. In explaining the elements comprising Figure 2.2, *The rehearsing and performing cycle* refers to the bands and their members involved in the Performance Management module. For the *Students rehearsing* activity of the cycle, I decided to explore a number of approaches to *Intra-peer assessing.* These are represented by *Exploring intra-peer assessment* in which identifying the contribution of each student to rehearsing was of interest to me. I began to explore students' personal attributes that they considered relevant to rehearsing, and how peer assessment might develop these attributes.

The *Students performing* element, relating to bands performing on stage, involved summative assessment. I formed the student *Inter-peer assessment panels* that assessed these band performances. Their performances were additionally marked by myself and sometimes by other tutors. *Students' Responses* and *Exploring peer assessment processes* also describe the quantitative assessment data arising from these performances. During each cycle, I would have the opportunity, explained in *Exploring the peer assessment process*, to intervene and change particular aspects of peer assessment if I considered it appropriate. This might include altering the timing of intra-peer assessments and whether or not to allow these to be conducted privately or

collaboratively among the band members. In deciding on an intervention, I thought it important to establish whether this might disadvantage students, whether as a band, or as an individual. There might be many occasions, however, where I decided not to change a particular aspect, preferring instead to repeat the cycle as before, in order to clarify, compare, or confirm any effects. On these occasions, any interventions would be informed by my increasing experiential knowledge of peer assessment. Interventions would also be guided by my values, explained in Section 2.8. *Students' responses* would arise from the quantitative assessment data and from interviews with band members.

The conclusion of each cycle was a natural place where I could consider what I had learned. As indicated in Figure 2.2, this would involve reflecting on the following: *Development of my understanding of peer assessment; Development of the research investigation*; *Development of the methodological framework.* Finally a *Revision and re-application of the action research cycle* restarts the next loop.

Of course, there would be occasions during the cycles where it was appropriate to reflect on students' responses and the process aspects before completion of the cycle. The end of each cycle, however, generally coincided with a short break for students (an end of term break, or a private study week), which provided me with a little time in which to reflect on the next action research cycle.

2.7 Role and ethnographic considerations

A number of paradigmatic assumptions were considered in Section 2.2.1. These involved assumptions made about my research and tutor roles being inseparable from the observed students and their bands in activities whereby knowledge might be constructed. Although such activities, including my interaction with the students, would occur within a natural educational and musical situation, I considered, nevertheless, that my participation as a tutor-researcher would inevitably have some impact on this setting.

The purpose of this section, therefore, is to examine my role as tutor among the community of learners and, my role as researcher in the same. I shall firstly consider the impact I might have as a tutor upon my research, and secondly, the impact I might have as a researcher upon my role as a tutor.

2.7.1 Impact as a tutor upon the research role

Being a tutor to the students placed me in a position of controlling certain aspects of the action research cycles of rehearsing and performing; for example, the timing and sequence of the various processes of peer assessment. I would also be able to determine the group membership; deciding whether to maintain existing group formations, change these, or allow students to select these by themselves. In addition, I would facilitate the formation of the inter-peer assessment panels, drawn from third year performance students[3]. The responsibility of establishing criteria for inter-peer and intra-peer assessment would similarly be mine; these activities are explained in Section 3.3.1. Other examples of tutor control impacting on the research would include: devising the project briefs for each assignment; determining performance contexts, such as that of the venue; deciding the method of assessment.

Although I was the sole tutor involved in the *Performance Techniques* and *Performance Studies* classes, the delivery of the *Performance Management* module, however, was not entirely my responsibility. During 2000-2003 another tutor would, on occasions, team-teach with myself and attend the performances[4].

It was from these aspects of my role as a tutor: determining groups; responsibility for the criteria; devising the assignments, which I considered had a particular impact on the research. I shall now examine the opposite: the impact of the research on my tutorial role.

2.7.2 Impact as a researcher on the tutor role

The issue, raised earlier, that my involvement as a researcher might disturb the natural setting of band rehearsal and performance, required careful consideration. In order to minimise this disturbance, I believed that adopting a research role that would be as unobtrusive as possible, would lessen its impact on the natural learning and teaching environment. There would be a number of unavoidable intrusions nevertheless.

---

[3] Acting as an interpeer panel assessor would be on a voluntary basis, and this aspect is discussed further in Section 3.1.1.1.

[4] My colleague also double-marked the performances and provided occasional feedback to the bands on their performances, but did not contribute, advise or assist in the research.

Examples of these might include: explaining and discussing the research study with the class and, if they agreed, their involvement in it; conducting interviews with individuals or a band.

Also, by adopting an action research approach, it would be probable that my researcher role would have a continual presence in many of the learning and teaching activities. For example, it would impact on the regular pattern and sequence of the weekly Performance Management classes as a consequence of the various interventions in the rehearsing and performing cycle. I would need to be alert to the possibility that my interventions might adversely affect the continuity or progression within the Performance Management course. I was also aware that the interviewees might regard the interviews I would conduct as a research activity that was separate from their normal educational experience

2.7.3 Ethnographic considerations

The previous discussion considered the likely impact that my researcher and tutor roles would have on each other within the natural setting of the enquiry. The purpose of this section is to identify the possible ethnographic characteristics of its approach. Firstly, this research topic can be described as one that engages in practical activities (the rehearsing and performing cycles), practical circumstances (the educational context and naturalistic setting) and social interaction (investigating how peer assessment might assist students' learning in group work activities). Secondly, for social constructivism, the creation of knowledge in the context of rehearsing and performing involves my entrance into, and interaction within, the life-world of the bands. In bringing together these observations, the research might also be likened to that of an ethnographic enquiry, following the account of Atkinson & Hammersley (1994):

> Central to the way in which ethnographers think about human social action is the idea that people *construct* the social world, both through their interpretations of it and through the actions based upon those interpretations. (p.27).

This enquiry is concerned with students in their natural settings in order to acquire a deep understanding of their involvement in peer assessment. Ethnography is a long process, however, requiring the researcher to be immersed in one social setting for an extended period of time (Bryman, 2004). Although not a pure ethnography, this study

nevertheless, has an ethnographic approach in view of its extended timescale, frequency of my visits to the settings and sustained period of data gathering (Murtagh, 2007). Furthermore, the data and its analysis would be inductive rather than deductive. Hypotheses, as such, would begin to emerge *in situ* rather than *a priori*; mixed methods of data collection and analysis, combining quantitative and qualitative data, could be likened to the colourful description of the ethnographer being a 'methodological omnivore' (LeCompte & Preissle, 1993, p.232).

In summary, a number of issues involving the impact of my tutor and research roles upon the enquiry were examined. It identified a number of features that, as a tutor, I would be able to control. It also considered the impact on my role as a tutor that the research would have, such as interventions in the rehearsing and performing cycle and conducting interviews. Finally, as a consequence of my participation as a tutor and researcher, the topic, its naturalistic setting and my involvement within the life-world of the student bands, the enquiry would have certain characteristics in common with the 'quasi-ethnographic approach' described by Murtagh (p.193).

2.8 Consideration of values

This section considers the impact of personal values on the enquiry. Action research differs from other methods because of its focus upon the researcher's values; these can be as important as methodological factors (McNiff *et al.*, p 14). Adopting an action research approach, therefore, required me to think carefully about the impact that my personal values might have on this enquiry. Action researchers can be described as 'working intentionally from deep-seated values that motivate them to intervene' (*Ibid,* p.10). If so, then a basis for 'interventions' in the rehearsing and performing cycles, for example, would be that of bringing about a situation that was congruent with my values. At the centre of these was a belief in the potential of peer assessment to improve student learning.

I considered that this enquiry and its methodology would be informed by personal values of a moral character: respect and belief; trust and faith; risk-taking. Indeed, as Aspin (1984) asserts, the principal justification for engagement in education can only be conceived in terms of its moral purposes.

The students' commitment towards and love of performing their music deserves respect; popular music, historically and sociologically, has always been considered as an intrinsic part of youth culture. It is for reasons such as these that the curriculum domain of popular music can be thought of as being especially identified with the life-world of the student community. I particularly value their knowledge of, and insights into, contemporary styles and genres. Their enthusiasm and desire to share their musical tastes with others, including myself, especially of contemporary idioms in which I might not always be as assured as I would wish, commands respect. Arising from these considerations is a belief that popular music students have particular potential to become effective peer assessors and provide insightful feedback, as well as learning from the experience themselves.

Being regularly and often closely involved with students during their group rehearsals has, over time, enabled me to understand the importance of trust in my professional practice. The importance of trust for this enquiry, for example, would be the nature of the student-tutor relationship, the peer-to-peer student relationship and also the student-peer group relationship. Having faith in students to demonstrate their responsibilities through marking their peers fairly, I hoped also, that they might have faith in my ability to ensure that the peer assessment process would be conducted properly. Finally, I would trust undergraduates to display their maturity, in allowing their personal rehearsal qualities to be involved in intra-peer assessment. Trust would be important for the study; I believed that, given sufficient trust among those involved, this would be a meaningful enquiry.

Providing sufficient trust existed within and between students and myself, there would be opportunities for taking certain risks in the interventions, where these were justifiable. I have always held a positive attitude towards thoughtful risk taking, either in musical performance or in certain aspects of learning and teaching. Risk-taking might, for this enquiry, encompass a range of situations in seeking solutions to problems encountered in the peer assessment of rehearsing and performance.

The dangers of taking risks in peer assessment and the possible damages incurred as a consequence of a poorly judged risk would need to be properly considered, however.

Among those consequences might be the following: animosity between individuals; apprehension about the usage of personal qualities as peer assessment criteria; student hostility to peer assessment; dysfunctional bands arising from biased assessments; deteriorating student relationships with the tutor.

Nevertheless, when contemplating the taking of a risk, I would need to make a judgement about whether this might disadvantage students, either as individuals or collectively as bands. The basis of my judgement would depend on whether I could reasonably foresee bands or their members being disadvantaged. Indeed, if carefully judged risk-taking contributed to the advancement of educational knowledge for all, then a justification, argued on moral grounds, might be offered.

2.9 Supporting learning

The central concern of this thesis is that of supporting learning and how peer assessment may assist this. I consider that peer assessment offers much potential for supporting students when engaging with certain activities that enable particular learning opportunities. Many of these were identified in Chapter One: formulating assessment criteria; providing cooperative learning and judgemental experiences in peer panels; encouraging students to devise and reflect on peer feedback. Table 1.1 (p.64), especially, suggests a range of areas in which this enquiry might support key learning experiences.

2.10 The curriculum area

Popular music is an area within the curriculum that, for many reasons, seems particularly suited to group work activities. For students of popular music and professional musicians alike, working in a group is a natural activity occurring in many contexts: performing, recording, production, sound engineering and song- writing are typical collaborative and creative activities involving musicians working together. As a musician for over twenty-five years and having experienced working in all of these contexts, I have learned much about the kinds of interpersonal skills that are of value in group-based musical creativity. I have an interest particularly in rehearsing and performing and these activities in a popular music setting can be quite different from

those typical in western classical music traditions. These differences can have implications for the kinds of interpersonal skills and interactions occurring within each. For example, rehearsing and performing classical music typically involves reference to a score; this may limit the scope for individual performance deviation, when compared with the liberties that are apparent in popular music performance. Rehearsal and performance conventions in popular music depend much less upon the use of a score. Even where musical notation is used, abbreviations and approximations are typical (chord symbols, chord charts and basic tablature, for example) and these can provide performers with considerable latitude for interpretative freedom. Band rehearsing, in popular music, can be a highly collaborative activity; the absence of a score, or the performance convention of a conductor, typically involves an organic and evolving musical creativity that is dependent upon the contributions of each member of the band. Individuals' contributions themselves will, in part, depend upon the interpersonal skills and attributes of each participant. Rehearsing and teamwork skills are qualities that are highly prized among professional musicians; this research explores how peer assessment might assist the development of these.

2.11 Summary

This chapter considered themes and issues thought to be critical to the methodology of the thesis. It presented a context to the enquiry involving the knowledge domain of popular music, its setting, its participants, its assumptions, its emerging design and values that were implicit. Among the key themes and issues to arise were the following: constructivism; rehearsing and performing cycles; action research and research 'interventions'; personal values including trust, respect and attitudes to risk-taking; supporting students' learning in the popular music curriculum.

The assumptions, themes and issues examined within this chapter are, I consider, crucial to the development of an appropriate methodology for the enquiry. This exposition, therefore, is an important contribution to the foundations, which will underpin Chapter Three: *Methods of Enquiry* and indeed the thesis as a whole.

Chapter Three: Methods of Enquiry

3.1 Course, module details and students

3.2 Commencing action research

3.3 Developing the intra-peer assessment enquiry

3.4 Developing the inter-peer assessment enquiry

3.5 Developing the process features enquiry

3.6 Data collection and methods of analysis

3.7 Standards of quality and verification

3: Methods of Enquiry

This chapter explains the methods used to investigate peer assessment in the context of this study. The discussions contained within Chapter One: *Literature Review* and Chapter Two: *Methodological Considerations* have informed these methods of enquiry and reference is made to these, where appropriate. Following a brief overview of the enquiry, this chapter is organised into the following sections:

3.1 Course, module details and students
3.2 Commencing action research
3.3 Developing the intra-peer assessment enquiry
3.4 Developing the inter-peer assessment enquiry
3.5 Developing the process features enquiry
3.6 Data collection and methods of analysis
3.7 Standards of quality and verification

A background to this study of popular music group performance and peer assessment was provided in the *Introduction* and Chapter Two *Methodological Considerations*. It was explained that it was a naturalistic enquiry, located in particular contexts: a curriculum context (undergraduate popular music group performance), a location context (the rehearsing and performance activity taking place in the students' natural setting) and an assessment context (intra-peer and inter-peer assessment). The enquiry has a fundamental focus on educational practice involving undergraduates rehearsing in their bands followed by a performance. As this pattern of activity was repeated a number of times during the year, it was described, in Section 2.4, as a 'rehearsing and performing cycle'.

Three lines of enquiry were established: inter-peer assessment, intra-peer assessment and process features of peer assessment. The following is a summary of the activities and the issues that were explored in each of these.

(i) The inter-peer assessment enquiry investigated what students on a peer panel might learn during their inter-peer assessing experiences. The enquiry also compared the marks awarded by peer panels with those awarded by tutors.

(ii) The intra-peer assessment enquiry involved investigating the use of students' personal attributes, in the context of band rehearsals, as assessment criteria.

(iii) The process features enquiry involved an investigation into how peer assessment processes impact on students. Such processes might, for example, include: the timing or sequence of peer assessment activities; conducting the marking in private or in collaboration; deciding whether marks should be awarded for formative or summative purposes.

Common to each of these lines of enquiry was an exploration into how learning might be supported and developed through the use of using peer assessment in popular music group rehearsing and performance.

3.1 Course, module details and students involved in the research

Contextual details about each of the courses and their contributing modules (*Performance Management, Performance Techniques 3* and *Performance Studies*) were explained in Section 2.1 (pp.73-75). Unless otherwise indicated, the use of the word tutor in this chapter refers to myself.

BA (Hons) Popular Music Studies Performance Management Module
The majority of students matriculated with non-traditional qualifications, typically a National Diploma in Popular Music. Many of these had performing experience in bands at their previous FE colleges as well as in bands unconnected with their studies.
A total of five first year cohorts (from 2000 until 2005) of Performance Management students, were involved in the research, comprising 170 students. Out of this total, approximately 50 were mature students, many of whom had substantial professional and semi-professional gigging experience involving a number of bands. Differences of age between students were not an issue; indeed there was mutual respect arising from the shared experiences of musicians from a variety of musical backgrounds. There was a gender imbalance however, typical of popular music courses, with males comprising 121 out of 170 students. Instrumental expertise among individual male students, ranked in order of frequency was as follows: guitarists; electric bassists; vocalists; keyboard players; and drummers. The overwhelming majority of females were vocalists. Most students were proficient on other instruments: many guitarists played bass and drums; many vocalists played guitar or keyboard. Irrespective of what was considered as their first instrument, at least half of the 170 students also contributed vocally, as lead or backing singers.

Three of the four assignments conducted during each year involved peer assessment. Students received a project brief for each assignment, which described its aims, learning outcomes and assessment details.

The aims and learning outcomes of each assignment remained unchanged over this period. The following two points summarise the information that appeared on all the project briefs during the five-year period, examples of which appear in Appendix 1 (p.271) and Appendix 2 (p.273).

- The aims and learning outcomes of the project were described, together with the context of the performance and, how the group membership was to be selected. The assessment criteria for the performance were described (if these were imposed by the tutor) or, if students themselves were to generate their own assessment criteria, this was explained, and then subsequently distributed.

- The assessment details and marking system used for the assignment also appeared on the brief; they specified the date of the completion of the assignment (usually the performance date) and the date when the marks and feedback would be returned to the student.

Many learning outcomes of this module were related to the acquisition of skills, knowledge and understanding that related to group rehearsing and performance activities. The module specification also allowed tutors to decide on the method used for determining band membership for each assignment.

Tutors themselves decided group membership for the first assignment, which was also tutor-assessed. The cohort was organised into a number of groups, each group comprising a typical band line-up of a singer, a guitarist, an electric bass player and drummer. Quite often, an extra guitarist, vocalist or keyboard player was included, depending on the range of instrumental expertise that was available within each particular cohort.

It was usual for the tutor to maintain the same band membership for the second assignment, although slight changes to groupings were necessary if, for example, a student left the course. For the third and fourth assignments, however, students were normally al-

lowed to select their own groups, with the proviso that the tutor reserved the right to alter the groupings where necessary.

BA Extended Degree Performance Studies Module

Performance Studies was a module belonging to a one-year HE foundation course, successful completion of which allowed direct entry on to BA (Hons) Popular Music Studies. Over half of the total of 43 who were involved in this study were mature students and the range of instrumental expertise was comparable with that of Performance Management. Again, there was a gender imbalance: 39 male and 4 female students. The Performance Studies module was similar to the BA1 Performance Management in its aims, learning outcomes and group work activities. As explained later in this section, the module was involved in the research for two academic years only, as a consequence of a refocus of the enquiry.

BA (Hons) Popular Music Studies Performance Techniques Year Three Module

Performance Techniques was an optional Honours level module involving solo and band activities that attracted students with advanced performance skills. As explained in Section 3.1, students were formed into panels in order to inter-peer assess the performances of bands belonging to the *Performance Management* and *Performance Studies* modules. Almost all of these students had previously undertaken the Performance Management module during their first year. Student profiles were comparable with the other modules: guitarists and vocalists were predominant and their ages ranged from twenty to forty years. *Performance Techniques Year Three*, along with that of *Performance Studies*, was also involved in the refocus of the enquiry.

In addition to friendship groupings occurring within each module, this occurred also among students who were performers on the same instrument. There were also a few friendship groupings apparent between courses and years.

Institutional permission for conducting this research was given by the Faculty of Higher Education at Barnsley College. The Head of Music at Barnsley College agreed to the particular modules, classes and the activities that would be involved in the enquiry. At the start of each academic year, each cohort of students was informed about the research into the use of peer assessment. It was made clear that participation in the research was voluntary. A presentation was given to each class in which it was explained that in-

volvement in the peer assessment activities would be subject to their agreement. The outline of the presentation, in which the principle of agreement, transparency, training and the rationale for peer assessment are described, appears in Sections 3.3 and 3.4. Many third year Performance Techniques students who would be participating in peer panels were already aware of the research into peer assessment from their own involvement, two years earlier, in Performance Management.

A review of the research progress was conducted at the conclusion of the 2001-2002 academic year. Rather than attempting to sustain the investigation across all three lines of enquiry (i.e. intra-peer assessment, inter-peer assessment and process features), a refocus might enable key areas of the research, which appeared to be more promising towards supporting learning, to be explored in greater depth. One of these areas involved the intra-peer assessment enquiry with its exploration of personal attributes as criteria for assessment. Similarly, the process features enquiry was already proving useful in guiding the intra- and inter-peer assessment activities. I concluded therefore, that concentrating the investigation on these areas might allow a deeper and more satisfying research enquiry. As a consequence, I decided to discontinue the inter-peer assessment enquiry, although the marking data arising from the peer panel assessments would continue to be gathered.

In order to enable a more concentrated investigation of band members' responses to intra-peer assessment, I decided to focus on the Performance Management module only. The Extended Degree Performance Studies module was more marginal to the enquiry, in terms of opportunities for rehearsing and performing and, in view of this, I decided to discontinue its involvement in the research.

From this point onwards, the term *Phase One* refers to the investigation that was conducted during the period between September 2000 and August 2002. *Phase Two* refers to the investigation following the refocus, from September 2002 onwards. A summary of the courses and modules involved in the research, the research focus, together with the numbers of students in each cohort, appears in Table 3.1 (p.103).

Table 3.1 Summary of the courses and modules, research focus and student cohorts

Year/Cohort		BA1	Extended Degree	BA3
Module		Performance Management	Performance Studies	Performance Techniques 3
Research focus		• Rehearsing and performing • Intra-peer assessment	• Rehearsing and performing • Intra-peer assessment	• Peer Panels • Inter-peer assessment
2000-2001	Phase One	59 students	25 students	30 students
2001-2002		41 students	18 students	25 students
2002-2003	Phase Two	34 students	Discontinued	34 students
2003-2004		21 students	Discontinued	11 students
2004-2005		15 students	Discontinued	10 students
Student totals		170 students	43 students	110 students

3.2 Commencing action research

The appropriateness of using an action research methodology for this enquiry was considered in the previous chapter, particularly in Section 2.6. Apart from its distinctive focus on practice, an action research approach seemed wholly appropriate to this kind of enquiry. Action research can have a flexibility that allows the researcher to draw upon a variety of methods pertinent to the enquiry. The flexibility of action research might, for example, enable an enquiry involving both quantitative and qualitative data. Using a mixed methods approach was considered in Section 2.5 and will be discussed again in Section 3.2.1.

Typical among action research studies are activities that can be described as an *action research cycle*. The *action research cycle* in this study was explained in terms of band rehearsing and performing, discussed in Section 2.4. The relationship between the action research cycles and the course rehearsing and performing cycle was indicated also in Figure.2.1 (p.71). During this enquiry, there was a total of some nineteen rehearsing and performing cycles. For Performance Management, each cycle was described by its assignment title and this structure was repeated each year: *Christmas Party; Venues & Audiences; Decade tribute*. Similarly, for Performance Studies, the name of each cycle was related to its assignment and a summary of each module appears in Table 3.2 (p.104).

Table 3.2 Group rehearsing and performing cycles

Cycle	Rehearsing & Performing	Course	Assignment title
Phase one			
1	Nov-Dec 2000	BA1	Christmas Party
2	Nov-Feb 2001	ED	Themed Set
3	Jan-Mar 2001	BA1	Venues & Audiences
4	Mar-Apr 2001	ED	Vocal Harmony
5	Apr-May 2001	BA1	Decade tribute
6	May-June2001	ED	Artist Tribute
7	Nov-Dec2001	BA1	Christmas Party
8	Jan-Mar 2002	BA1	Venues & Audiences
9	Apr-June 2002	BA1	Decade tribute
10	May-June 2002	ED	Artist tribute
Phase two			
11	Nov-Dec2002	BA1	Christmas Party
12	Jan-Mar 2003	BA1	Venues & Audiences
13	Apr-May 2003	BA1	Decade tribute
14	Nov-Dec 2003	BA1	Christmas Party
15	Jan-Mar 2004	BA1	Venues & Audiences
16	Apr-May 2004	BA1	Decade tribute
17	Nov-Dec 2004	BA1	Christmas Party
18	Jan-Mar 2005	BA1	Venues & Audiences
19	Apr-May 2005	BA1	Decade tribute

The naturalistic setting of the enquiry, being in the students' and tutors' usual teaching and learning environment, also supported a reflection on educational practice. Reflection, acting on this, evaluating what might happen following a change in the cycle, planning further action and repeating the cycle were key activities for the tutor. Robson (2002) would regard all of these as typifying an action research project.

3.2.1 Mixed Methods

As noted previously, the enquiry suggested the involvement of data that was both qualitative (responses by students) and quantitative (marking assessment data).
Rather than focusing on a single specific over-arching question, the nature of this research involved a number of lines of enquiry in which the collection of both quantitative and qualitative data was appropriate. It was also envisaged that a methodology that allowed a flexible and responsive approach to data collection that might be conducted at different stages of the inquiry would be required.

Robson (2002) observes that pre-determined designs or a single method typically becomes focused on the researcher's perspective, while more flexible designs can follow the participants' perspectives. An approach that combines these can allow for both aspects to be studied. It was decided, therefore, that it was appropriate for this enquiry to combine qualitative and quantitative approaches and methods. These approaches and methods are described in Sections 3.3 and 3.4.

3.2.2 Deciding upon interventions

A discussion of the methodological considerations relating to interventions in the rehearsing and performing cycles appears in Sections 2.5 and 2.7. Within the model of action research formulated for this thesis (see Fig 2.1, p.71), the term *intervention* is used to indicate a change in an aspect of peer assessment taking place during a rehearsing and performing cycle by the tutor as a part of this research enquiry.

Decisions that were made involving why, when or how to intervene in a rehearsing and performing cycle were based a number of factors. Among these were the researcher's values (considered in Section 2.8), the researcher's experience in using peer assessment (considered in Section 2.3) and reflection on the responses obtained from the preceding cycles. One example of an intervention involved a change to the intra-peer assessment marking scale, from a four- to a five-point for the next rehearsing and performing cycle. Another involved a change to the sequence of assessment activities within the cycle by, for example, conducting the intra-peer assessment *before* bands had received their performance marks, instead of after.

No evaluative checklists or observational recording methods were employed when deciding to intervene as it was considered that the use of these instruments could be intrusive to the students' natural setting. Indeed, deciding against employing observational recording methods in this context is not unusual; action research is different from other traditions because of its focus upon the researcher's values, which are considered to be as important as methodological factors (McNiff *et al.*, 1996; McNiff & Whitehead, 2006).

All the interventions that were conducted within each of the nineteen rehearsing and performing cycles appear in the illustrations contained in Chapter Four: *Rehearsing and performing cycles.*

3.3 Developing the intra-peer assessment enquiry

The description of the intra-peer assessment enquiry is organised as follows.

3.3.1 The intra-peer assessment enquiry: phase one

3.3.1.1 Initial training activities in peer assessment

3.3.1.2 Training procedures for formulating students' personal attributes

3.3.2 The intra-peer assessment enquiry: phase two

The research, involving students belonging to the Performance Management module, commenced during the academic year 2000. It was decided to introduce intra-peer assessment into the second rehearsing and performing cycle, principally for the purposes of determining individual marks. A single holistic criterion *contribution to rehearsing* was initially employed. The 'zero-sum' method[1] was used and an evaluation sheet adapted from Hunter (1999) was given to each student asking him or her to rate each band member's contribution to the assignment. The scores were then totalled and expressed as a proportion of the performance mark awarded to the whole band. Figure 3.1 (p.107) is an example of the evaluation form that was used by each student during the first six cycles. The process by which the scores were totalled and how individual marks were calculated appears in Section 2 of the Appendix.

[1] The 'zero-sum' method was explained in Chapter One, Section 1.4.4

Did everyone contribute equally? If so, each member of the group should be given 25 % if the band is a four-piece (or 20% if the band is a five-piece, etc). If two members of the group contributed significantly more than their colleagues the distribution of % weighting might reflect this in the following allocation: 15+15+35+35. Allocate a mark for each member of the band. Ensure that marks together total 100.

Group members	Assessment	Group members	Assessment
1(self)_____	_____	4_____	_____
2_____	_____	5_____	_____
3_____	_____	Ensure marks total=100	

Explain and justify your mark for each member:

1_____

2_____

3_____

4_____

5_____

Report completed by_____ Course_____ Date_____

Figure 3.1 An intra-peer assessment form based on a general contribution to rehearsing criterion (adapted from Hunter, 1999.

The next section describes the use of students' personal attributes in rehearsing as intra-peer assessment criteria.

3.3.1.1 Initial training activities in peer assessment

Two short training sessions, conducted on separate occasions and undertaken during normal class time, were given to students prior to being involved in peer assessment. The first session involved an explanation of the rationale for using peer assessment including its activities; the second session discussed what might constitute *contribution* in band rehearsals.

The central aim of the intra-peer assessment enquiry was investigating its potential for developing students' learning. As described in Section 2.2.2, exploring the use of intra-peer assessment criteria based upon students' personal qualities or attributes became a key aspect of the research. Students' personal qualities or personal attributes, related to the context of rehearsing and performing in a group, could be considered as rehearsal-related learning criteria. I considered important for students to improve their under-

standing of personal attributes in this learning context. Identifying personal attributes for example, could also be regarded as a form of self-assessment, which could incorporate a mechanism for students to award marks to themselves.[2] The particular relationship between self-assessment and peer assessment, in the literature on peer assessment, was examined in Section 1.2.1. For the eighth rehearsing and performing cycle (Autumn 2001), students were asked to develop and establish a set of personal attributes that they considered *important* to their group rehearsing. In generating these, students were also being encouraged to take ownership of the assessment criteria.

3.3.1.2 Training procedures for formulating students' personal attributes

The previous initial training procedure for intra-peer assessment was modified and extended. It usually required between two to three hours and was divided into three parts: purpose and process of intra-peer assessment; using personal attributes as assessment criteria; identifying students' own personal attributes.

Training commenced with a discussion of the assignment brief and an explanation of the rationale and activities involved in peer assessment. This was followed by a session that introduced the concept of personal attributes (using the terms 'attributes' and 'qualities' interchangeably) in the context of group rehearsing.

A class-based activity followed, which encouraged students to suggest personal attributes that they thought might be appropriate for use as intrapeer assessment criteria. It began with a brainstorming session in which students were asked to identify personal attributes arising from their experiences of rehearsing in bands that they considered should be displayed in their rehearsals. These rehearsal-related personal attributes were compiled on a white display board. The class further considered and discussed each attribute in terms of their possible meanings for the context of band rehearsing. Students, working individually, were then invited to consider five or six of these attributes (or any others that did not appear during brainstorming) that they regarded, arising from their own experience, as being the most important for them. The purpose of this was that of developing students' awareness of their personal attributes and being able to articulate those that they considered important for their rehearsing.

---

[2] The links being made here between students' personal attributes, learning criteria and assessment criteria recalls Norton's (2004) paper on using assessment criteria as learning criteria, discussed in the Literature Review (p.37).

The next step involved exploring students' personal attributes in relation to two areas: attributes that students together in their bands decided were important to them as an entity; attributes those students, as individuals, regarded as being appropriate.

The nature of band rehearsing and performing being illustrative of an *in situ* collective learning experience was discussed in Chapter Two; this experience, I considered, should be reflected in criteria formulated by the band. Consequently, bands were asked to devise three mutually agreeable *group attributes* that they believed were important to their rehearsing together, which could be used as intra-peer assessment criteria. It is important to stress that these were group-agreed attributes, formulated on what they, as a band, decided applied to them specifically, rather than attributes of a more generic group work character such as 'teamwork' or 'communication'. The use of the term *group attributes* from this point onwards in this research, therefore, refers to 'group-agreed' attributes and to the explanation given here.

Similarly, in order to give a focus to their *individual* involvement in rehearsing, each student was asked to formulate three *individual attributes* that they regarded as being of personal importance, which could also be used as intra-peer assessment criteria.

To summarise, each student generated a total of six personal attributes, comprising three group-agreed and three individually generated attributes, which they agreed to be used as their assessment criteria. The literature on peer assessment criteria, examined in Section 1.4.2.4, indicated that six was a suitable number that would allow an appropriate range of qualities to be assessed, while remaining manageable within the assessment process.

In order to convey the character of the meanings, symbolised within each of their attributes, students were asked to formulate what they considered to be the opposites. This resulted in a set of bi-polar descriptors in which one of the poles represented the positive or desired attribute, while the other pole represented the opposite or antonym; for example, *Patience - Irritability.*

The marking system that I decided to employ was based on three principles: flexibility of marking criteria; differentiation and sensitive discrimination (between, for example,

Patience and *Irritability)*; transparency and fitness for purpose in terms of providing marking feedback.

It was decided that using a semantic differential type of rating scale was appropriate for assessments across the bi-polar descriptors, as illustrated in Figure 3.2.

Patience $\underset{5}{\_}\ \underset{4}{\_}\ \underset{3}{\_}\ \underset{2}{\_}\ \underset{1}{\_}$ Irritability

Figure 3.2 Semantic differential rating scale based upon bi-polar personal attributes

There are cautionary factors about using rating scales: for example, we are reminded by Cohen *et al.* (2000) that there is no assumption of equal intervals between the categories. Hence although the space between *Patience* and *Irritability* that denotes a rating of 2, it does not necessarily indicate that this intensity of feeling is exactly twice as strongly felt as that denoted by a rating of 4. Similarly, one cannot infer that the rating of 2 across the *Patience - Irritability* rating scale necessarily corresponds to an equivalent rating of 2 across the *Contributing to ideas – Not contributing to ideas* scale. It was thought reasonable, however, to assume that each individual would be relatively consistent in applying his or her own marking judgements, based upon his or her own understanding of the criteria and the context of group rehearsing.

Evaluation, potency and activity measurements, often associated with the semantic differential, were not appropriate for this research. A five-point directional scale (*one* to *five*) was used, with *one* being described as *poor/negative*, *five* described as *excellent/positive*, and *three* described as *average/neutral*, with *four* and *two* being *above* or *below* average.

The literature on peer assessment indicated a preference for smaller scales (i.e. five-point scales). As it was anticipated that students would be apprehensive about assessing peers on the basis of their personal attribute criteria, I considered it unreasonable to expect individuals to discriminate across larger scales.

The sets of personal attributes, provided by each student, were then arranged into an assessment table and distributed amongst the relevant band members. The principle of seeking student agreement to each stage in the process was established.

3.3.2 The intra-peer assessment enquiry: phase two

Personal attributes in phase one (whether *group attributes* generated by the band, or *individual attribute* formulated by each individual), were focused on what students considered were *important* to them. In exploring how intra-peer assessment might help to develop students' personal attributes in their rehearsing, the basis on which *individual attributes* were determined was changed from *important* to those of *personal strengths* and *personal weaknesses*.

Asking each student to reveal what they regarded were their personal strengths and personal weaknesses required care and sensitivity, particularly regarding personal weaknesses. Encouraging students to be honest about their personal weaknesses and to be willing about sharing these with others, as intra-peer assessment criteria were delicate subjects that required careful consideration. Using students' personal attributes based upon their disclosures of 'weaknesses' could be problematic and pose considerable risk. Given that the principle of advancing learning was at the centre of this enquiry, taking risks in order to achieve this prize was, I considered, justifiable.
Employing students' *personal weaknesses,* as intra-peer assessment criteria, became an important and distinctive aspect of the enquiry.

A purpose for inviting students to consider also the qualities that they regarded as their personal strengths involved providing a balance. Employing personal weakness attributes alone might have resulted in an undue emphasis on negativity or, perhaps, raising students' anxieties about the process unnecessarily. These changes to the formulation of personal attribute criteria were explained and discussed with the class. As in their previous training sessions, student concerns or anxieties were addressed and agreement with them was sought where appropriate.

This process, therefore, culminated in students' generating three individual attribute criteria comprising two personal strengths and one personal weakness. Various combinations of individual attributes (personal strengths and/or weaknesses) and group attributes were employed during the rehearsing and performing cycles during phase two; these appear in the data collection and interventions described in Chapter Four.

An example of an intra-peer assessment form that uses three *Group Attributes* and three *Individual Attributes* (based on two self-selected *personal strengths* criteria and one self-selected *personal weakness criterion)* appears in Figure 3.3 below.

Christmas Party Dec 2002 Band F	INTRA-PEER ASSESSMENTS											
	Group Attributes			Individual Attributes (2 self selected personal strengths and 1self selected personal weakness) n/a=not applicable								
1-5 marking scale 1= negative/ poor 3= neutral/ average 5= positive/ excellent	Reliability (Unreliability)	Creative input (No creative input)	Commitment (Poor commitment)	Performing Confident (Nervous)	Making decisions (Letting others do the deciding)	Excellent cooperative skills (Poor cooperation)	Excellent performing ability (Poor performing ability)	Remembering equipment (Not remembering equipment)	Organisation skills (Lack of)	Leadership (Poor leadership)	Excellent interpersonal skills (Poor interpersonal skills)	Punctuality (Not punctual)
SS				N/a	N/a	N/a	N/a		N/a		N/a	
QB						N/a	N/a	N/a	N/a	N/a	N/a	
LP				N/a	N/a			N/a	N/a	N/a		N/a
MJ					N/a	N/a	N/a	N/a	N/a			N/a
CR					N/a	N/a	N/a	N/a		N/a	N/a	

Figure 3.3 Example of an intra-peer assessment form using Group and Individual Attributes

The rehearsing and performing cycles from this point onwards became increasingly focused on using students' personal weakness criteria, where I considered the potential for supporting learning, in terms of bringing about improving attributes, might be greater. A further refinement to the process involved the personal weakness criteria for each individual being determined by band members, rather than by the student him/herself. This was a procedure that required an appreciation of the considerable risks and difficulties that might be encountered. I decided that it was sensible that students' written consent was obtained in advance of any band-determined personal weakness attributes being formulated as assessment criteria. A different focus on students' personal weaknesses took place during the concluding action research cycles. It involved an intervention in the number of intra-peer assessment criteria that were employed; three group attributes being replaced by a single, tutor-imposed *general contribution to rehearsing* criterion. This altered the weighting between group attributes and personal weaknesses from a 50:50 balance to that of a 25:75 emphasis on personal weaknesses. For cycles 17-19, the weighting of personal weakness criteria also increased to 100% resulting in group at-

tributes criteria being discarded completely. A summary of the personal attributes terms discussed in this section and definitions applying to the remainder of this thesis appear in Figure 3.4.

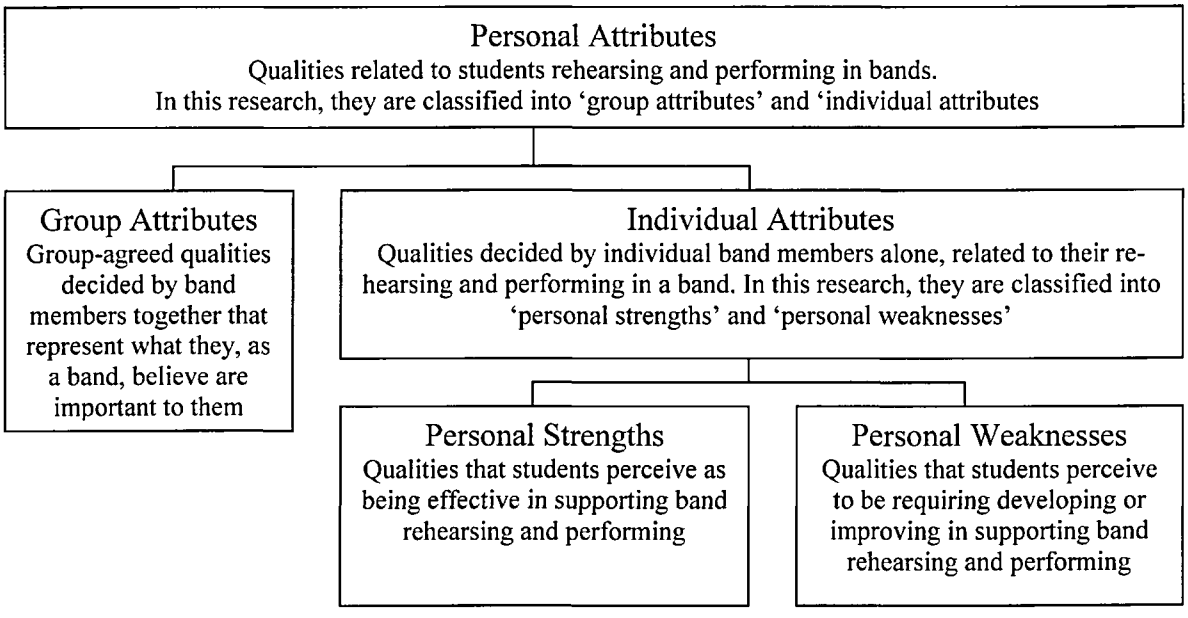


Figure 3.4 Summary of the 'personal attribute' terms used in this enquiry

3.4. Developing the inter-peer assessment enquiry

The description of the inter-peer assessment enquiry is organised as follows.

3.4.1 The inter-peer assessment enquiry: phase one

3.4.1.1 Peer panels and training

3.4.1.2 Development of inter-peer assessment criteria

3.4.1 The inter-peer assessment enquiry: phase one

The rationale for using peer panels drawn from the year three *Performance Techniques* class with which to assess the bands was explained in *Methodological Considerations*, in Section 2.1. The aim of the inter-peer assessment enquiry was twofold:

(a) exploring what peer panels themselves might learn during their inter-peer assessing experiences;

(b) comparing marks awarded by peer panel with those of the tutors.

3.4.1.1 Peer panels and training

It was decided to form peer panels comprising four students: this number would enable a mix of instrumental and performing expertise, addressing various issues raised in the literature (see Section 1.3.3.); it would permit a range of discussion about the band performances but also be manageable in resolving differences and reaching agreements.

Membership of peer panels was voluntary; participation in inter-peer assessment was not a requirement of the module. Forming peer panels on the basis of volunteers, however, could impact on the research if, for example, one volunteer had a distinctive, domineering or disproportionate input. In view of providing opportunities for students to develop their evaluative and judgemental skills through this assessment experience, it was intended that as many as possible would be offered the option of participating in a peer panel. Selecting panel members from the volunteers was, therefore, informed by the principle of inclusion: involving as many individuals as possible from the Performance Techniques class. Students who had yet to participate in a peer panel were favoured over those who had already acted as an assessor. A second principle, upon which volunteers would be selected, involved the provision of a range of instrumental expertise. A panel that comprised a guitarist, vocalist, drummer and keyboardist would be preferred over one that consisted of four bass players. It was hoped that the selection of volunteers on these two principles would help to minimise the disproportionate input of any one particular individual.

The peer panels were provided with training sessions in preparation for the assessment of the Performance Management bands. Their training included a multimedia presentation and discussion of the rationale in addition to mock assessment and marking sessions using video-recordings of previous band performances. Peer panel members were, however, familiar with evaluating undergraduate performances through attending a weekly performance workshop in which students performed to each other and received peer feedback; this is described in Davis & Pulman (2001).

3.4.1.2 Development of inter-peer assessment criteria

The Performance Management class generated the assessment criteria for the performances themselves. The method employed for establishing inter-peer assessment criteria

was similar to that used for intra-peer assessment criteria. It began with a brainstorming activity whereby individuals suggested assessment criteria for their band performances, which were written on a display board. An extensive list of possible criteria was typically generated and the class voted for the criteria that they favoured. Some criteria were discarded because of their similarity to others that had already been chosen; other criteria were discarded simply through being inappropriate.

The next step was that of deciding whether their performance criteria should be given category weightings, or be interpreted holistically. A class vote was conducted, wherever appropriate, at each stage in the process.

One or more tutors also separately assessed each performance and, where necessary, moderated the student marking. The possibility of tutor moderation of marks was made clear to the peer panels as well as to the Performance Management bands in their respective training sessions.

Sufficient time was given for the peer panels to reach their decision about the marks and nature of the written feedback following the band performances. When panels indicated that they had completed their assessments, the tutor(s) would then join the student assessors to receive their comments and the marks that they had agreed. Peer panel students understood, as part of the training sessions, that they were expected to observe the ethics and confidentiality of the activity. The marking procedures that were adopted were intended to minimise peer panels' exposure to attempts to influence their assessments from those whom they had assessed. The post-performance meeting and agreement of their marks with the tutors that was conducted immediately after the performances restricted opportunities for unethical practices. The inter-peer assessment enquiry was discontinued following the review of the research that took place at the end of phase one, as previously explained.

3.5. Developing the process features enquiry

The importance that was attached to acquiring competence in the operational mechanics of peer assessment, necessary to progress the enquiry, was considered in Chapter Two, Section 2.3. The primary objective of phase one in the process features enquiry was the achievement of this.

Process features of peer assessment naturally overlapped the other two lines of enquiry: for example, in the activities through which assessment criteria were generated. There were two important process features, however, that did not overlap with the investigations being conducted in either the intra- or inter-peer lines of enquiry. The first of these process features involved exploring students' responses to intra-peer assessment that was conducted in either a private or collaborative setting[3]. The second involved an enquiry about *when*, during the rehearsing and performing cycle, intra-peer assessment should be conducted.

Each of these process features and the responses arising from the action research were also investigated during phase two. Other areas of exploration included the provision of formative marking activities and self-assessments that would contribute to grades. Students' responses to these process features were examined in the quantitative assessment data and also during interviews.

3.6 Data collection and methods of analysis

The previous section described the three interrelated lines of enquiry and the development of their methodologies: intra-peer assessment; inter-peer assessment; process features of peer assessment. A description of the methods of data collection and analytical procedures now follows.

Mixed methods of data collection and analysis, involving qualitative interview data and quantitative assessment data were used; this was discussed in Section 2.5. A summary

[3] Sometimes, within this little explored area of intra-peer assessment in the literature, the term 'secret assessment' is used to denote peer marking that is conducted in private (for example, Lejk, M. and Wyvill, M. (2001b). The expression 'secret assessment' may be suggestive of something that is perhaps sinister. I prefer to describe such an activity as a 'private' assessment. A 'private' assessment setting also, I believe, emphasizes students' rights to, and respect for privacy. The term 'private' assessment, from this point onwards, is used to describe such an intra-peer assessment setting.

of the data collection methods and how each of these relates to the three lines of enquiries appears in Table 3.3.

Table 3.3 Summary of methods of enquiry, quantitative and qualitative data collection

Methods of Enquiry	Quantitative data collection	Qualitative data collection
Intra-peer assessment enquiry	Intra-peer assessment data	16 Interviews with individual band members
Inter-peer assessment enquiry	Inter-peer panel assessment data	3 Interviews with panel members
Process features enquiry	Intra-peer assessment data Inter-peer assessment data	16 Interviews with individual band members, as above

The description of the data collection and methods of analysis is divided into the following sections:

3.6.1 Qualitative data collection: interviews with individual band members
3.6.2 Selecting interviewees
3.6.3 Deciding on the method for analysis of the interviews
3.6.4 Qualitative data collection: interviews with panel members
3.6.5 Quantitative data collection: intra-peer assessment data
3.6.6 Deciding on the method for analysing the quantitative assessment data
3.6.7 Quantitative data collection: inter-peer assessment data
3.6.8 Qualitative and quantitative data collection and the process features enquiry

3.6.1 Qualitative data collection: interviews with individual band members

The qualitative data was gathered from individual interviews that were conducted with sixteen Performance Management students. Section 2.2.2 considers the paradigmatic assumptions that were made in relation to the use of interview data in this enquiry.

The overall purpose of the interview was to obtain students' responses to their involvement in peer assessment activities. I concluded that an interview design that would combine open-ended questions along with focussed lines of questioning, that were relevant to areas of the enquiry (for example, learning, process features, marking feedback), would be appropriate. Between the extremes of structured and unstructured interview techniques, is the semi-structured interview format.

Adopting a semi-structured format would permit a framework of questions, which allowed flexibility in their order, and appropriate follow-up questions that would enable time to explore in depth particular responses and issues that individuals raised. Studies influenced by constructivist thought frequently adopt semi-structured formats (Charmaz, 2000, 2006). A semi-structured interview framework, which was thought to satisfy both open questions and more focused questioning, was devised. Interviews began with open-ended, non-judgemental questions in order to encourage free responses to emerge. It then moved to an exploration of peer assessment topics, of particular relevance for this enquiry, guided by semi-structured focused questions. The interview concluded with other open-ended questions, in order to encourage students to comment on any aspect that they considered might improve the peer assessment experience. The final open-ended question invited the student to comment upon any aspect that s/he felt had not been explored in the interview. The interview framework of questions that was adopted for the Performance Management students was based on the following.

(a) Open questions inviting responses to the intra-peer assessment experience;
(b) Focused questioning involving personal attribute criteria;
(c) Focused questioning involving process of awarding marks in private and in collaborative settings;
(d) Focused questioning involving the timing of intra-peer assessment;
(e) Focused questioning involving intra-peer assessment feedback and marks;
(f) Focused questioning involving assessing band members;
(g) Open questions involving suggestions for improving the intra-peer assessment experience;
(h) Open questions inviting issues that had not been already explored.

This format allowed additional questions to be introduced or students' responses to be explored more deeply, when appropriate to the enquiry, using follow up questions. An example of the framework of questions that was adopted appears in Appendix 4 (p.278).

3.6.2 Selecting interviewees

The selection of interviewees was obviously restricted to the students who were involved throughout each *Performance Management* cohort. Interviews with students were selected on a voluntary basis and were conducted in the students' own time. The

invitation that was verbally issued to students for possible participation in interviews took place at the conclusion of the academic year. All of the students (except one) from each cohort signified their willingness to be interviewed. I considered that five interviews conducted from each year (2003-2005) would be invaluable in obtaining a range of responses to intra-peer assessment while remaining manageable in terms of time and resources. Volunteers were selected on the basis of mutual availability regarding the day or time of interview. The purpose of the interview, in contributing to the research, was made clear to potential interviewees.

Students were assured that individuals would not be named or identified in the transcript of the recorded interview, involvement in the research, or in subsequent publication. Furthermore, the research was governed by the research ethics regulations of the institution. There were sixteen interviews with individual students, each being invited to sign a written permission form, or signify their permission on the recording itself. No students over the duration of the study objected to their inclusion in the research.

Interviews were conducted in a small band rehearsal room, which would be very familiar to the students. The rehearsal room was conducive towards creating a relaxed atmosphere and was considered to be an intimate and friendly interviewing environment.

The interview was recorded onto a minidisk, audiocassette or burned directly onto CD, with the microphone placed centrally between the student and myself. Care was taken to ensure that the recording equipment was as unobtrusive as possible. Permission to record the interview was sought, with the guarantee that no one else (except my supervisory team) would hear the interview.

3.6.3 Deciding on the method for analysis of the interviews

Lincoln & Guba (1985), Denzin & Lincoln (1994), Miles & Huberman (1984, 1994) together with the literature surveys of Cresswell (1998) and Cohen *et al.* (2000) are among the key texts that tease out the characteristics and implications for data analysis within research conducted in natural settings, such as the context of this study. Typical of data derived from naturalistic enquiries is that the data suggests the theory rather than *vice versa*, that data analysis becomes an 'inductive process for illuminating social processes' (Miles & Huberman, 1984, p.20) rather than *a priori* and deductive, that the-

ory emerges rather than being pre-ordinate, and that *a priori* theory is replaced by grounded theory (Lincoln & Guba, 1985, pp.39-43; Cohen *et al.*, 2000, pp.137-138). I decided that a grounded theory approach would be an appropriate method for analysing and interpreting the interview data.

I utilised a number of grounded theory analytical procedures including those of open coding, axial coding and selective coding. The semi-structured interview data was analysed initially, using an *open coding* procedure where students' responses were placed into *conceptual sub-categories*. Connections were then developed between the conceptual sub-categories to form core categories, a process similar to *axial coding.*

The basis upon which sub-categories were identified was derived from the number of interviewees whose responses were represented by that category, together with the number of open coded responses attributed to it. I decided that where similar open coded responses arose from at least three of the sixteen interviewees (and also attracted several comments during their interviews) this constituted grounds for a conceptual sub-category. A decision on whether to differentiate between sub-categories further in terms of strength of responses, arose from the data analysis, rather than specified in advance. This data suggested two other measurement boundaries: one that contained the majority of sub-categories (i.e. those comprising similar responses from between 4 and 7 interviewees), the other containing the sub-categories comprising similar responses from between 8 and 16 interviewees). This is discussed further in the interview data analysis that appears in Section 5.1

The data, therefore, suggested three measurement boundaries:

(a) sub-categories defined by similar open coding responses from three separate interviewees that additionally attracted several comments during their interviews. The strength of expression for these sub-categories in the data analysis is described as 'weak' in the data analysis (p.164) of Chapter Five;
(b) sub-categories defined by similar open coding responses from between four and seven separate interviewees that additionally attracted several comments during their interviews. The strength of expression for these sub-categories in the data analysis is described as 'moderate' in the data analysis (p.164) of Chapter Five;

(c) sub-categories defined by similar open coding responses from between eight or more separate interviewees that additionally attracted several comments during their interviews. The strength of expression for these sub-categories in the data analysis is described as 'strong' in the data analysis (p.164) of Chapter Five.

The relationships between these core categories were then identified, integrated and synthesised diagrammatically. Developing this central integration focus was a procedure that could be likened to *selective coding*. Categories that indicated the central concerns arising from the interview data, whose potency reflected an active rather than passive character, or, categories that appeared to emphasise certain values, suggested a focus for integration. Theoretical saturation occurred with the final interview of the 2004-2005 cohort, although many individual categories had, by that time, become saturated. I considered that this grounded theory procedure that was adapted to the data, was fit for the purposes of this investigation. The analysis of the data and the procedures that were used is explained and discussed in Chapter Five.

3.6.4 Qualitative data collection: interviews with panel members

Two interviews were conducted with peer panel students and one peer panel group interview was also conducted. These three interviews took place at the end of phase one coinciding with my decision to discontinue the inter-peer assessment line of enquiry. As a consequence of this, no further peer panel interviewing was carried out.
As students worked together in panels on a common purpose (Watts & Ebbutt, 1987) with their experiences being regarded as intrinsically collective, a group interview approach was considered to be appropriate. Conducting a group interview also enabled the 'potential for discussions to develop, thus yielding a wide range of responses' (Cohen *et al.*, 2000, p.286).

The basis for interview selection was the same as for the individual Performance Management students, taking place in students' own time. Care was taken to present the interviews as being unconnected with students' academic studies, although it was suggested to interviewees that participation could be evidenced in their end of year portfolio.

At this early stage in the research, these interviews were also useful pilots for future peer assessment related interviews. Acquiring sufficient interviewing skills involving the particular topic of peer assessment also contributed to developing expertise in this sort of enquiry. The purpose of the interview was to explore peer panels members' responses to their experiences and what they may have learned through their participation. An unstructured interview method was adopted, as I considered open-ended questions were appropriate in eliciting students' responses.

3.6.5 Quantitative data collection: intra-peer assessment data

The kind of quantitative information that would inform the enquiry was carefully considered. As the volume of assessment marks generated during the period of the research would be considerable, this data was regarded as a potentially rich source of information. I considered that such data would be of particular relevance to the enquiry as indicators of the credibility in marking personal attribute criteria. These indicators would include, for example, measurements of: the consistency of peer agreement; the correlation of students' self-assessments with those of their band. This indicated four kinds of information.

(1) Information indicating the extent to which students agreed with each other when assessing their personal attributes.
(2) Information indicating the extent to which students under- or over-estimated themselves when self-assessing their personal attributes.
(3) Information indicating the proximity of students' self-assessment of their personal attributes with those of their peers.
(4) Information indicating the extent of agreement between band members.

Three data types were suggestive of the above: the marks that were awarded by the peer assessors; data resulting from statistical measurement techniques that were performed on those marks; data resulting from the correlation or interrelationship between these.

The first data type simply comprised the marks that were awarded and subsequently employed in calculating students' final grades for the module. This was described as *Primary Data;* it is explained further in the analysis that can be found in Chapter Five.

For the second data type, resulting from statistical measurement techniques that were applied to the extensive data arising from the 170 *Performance Management* students' marks comprising the intra-peer assessment forms of 75 bands[4], it was considered important to prevent a significant loss of detail. Interesting data about individual students, free riders and shooting stars, for example, might be submerged through an inappropriate use of high-powered tests or meta-analysis. I decided, therefore, that relatively simple, low-powered tests were preferable wherever possible. Simple tests, which, as far as possible, minimise the statistical manipulation of the students' peer assessment data, might help to maintain the connection between individual band members, bands, and the context of the assessments. Using sophisticated tests implies that the numerical data are accurate whereas the peer assessments obtained in this study can be considered at best an approximation. Sophisticated tests might give a spurious validity. Tests that would indicate patterns and distributions without suggesting high levels of accuracy would provide useful information. Indeed, for enquiries informed by constructivism, following Guba & Lincoln (1994), quantitative methods 'can play a useful informational role' (p.115). This has implications also, for deciding on an appropriate rationale that supports a method to analyse quantitative information that has not been subjected to high-powered statistical tests; Section 3.6.6 proposes this rationale. I considered that the following low-powered statistical tests, described below, offered simplicity, while maintaining a connection between individuals, bands and their peer assessment marking data appropriate to the spirit of this enquiry.

(a) Variance.

This is a simple measure of the difference between each of the band members' assessment of a student's attribute, and the student's own self-assessment, using the marking data. The mean of the band's assessment was compared with the student's self assessment resulting in a plus (+) variance score denoting a student's over-estimation of their personal attributes, and a minus (-) variance score denoting under-estimation. Variance was selected as a statistical measurement because of its simplicity and of it being grounded on the raw peer assessment data. Using the mean of the remaining band members' assessment in order to allow a comparison with an individual's self-assessment might conceal a wide divergence of group marks. Examining the standard deviation of group marks, however, would provide an indication of the spread of marks among the group.

[4] Further details can be found in Appendix 8 (p.291).

(b) Individual Normalised Difference Coefficient.

This straightforward measure was used to indicate the proximity of a student's self-assessment to that of the band members' assessment of that student. It was selected on a similar basis to that of variance; being a simple low-powered test requiring very little manipulation of the raw assessment data. Its purpose was to offer an alternative measure to that of variance. It can, for example, reveal the occurrence of a 'rogue' mark awarded by in individual band member. This test, however, indicates only the proximity of a student's self-assessment to that of their band members' assessment.

(c) Band Normalised Difference Coefficient.

This is simply the mean of each band's totalled individual normalised difference coefficients. As it is somewhat removed from engagement with the raw marking data itself, this is a measure of the proximity of the band members' self-assessments as a whole. It provides general information and patterns of data that rather than detail.

(d) Standard Deviation

Used for calculating the spread and divergence of numerical data, this was an appropriate statistical measurement to indicate the consistency of marking agreement involving intra-peer assessment of individual and group attributes. Mean standard deviations were also used in correlating consistency of agreements with other data. Although this test was of a higher power than the others described above, there was one unifying factor in that standard deviations were based, almost always, on a spread of four or five marks (i.e. four or five piece bands).

Marking data to which these statistical tests were applied were described as *Secondary data*. This is explained further in the analysis that can be found in Chapter Five.

The third data type, arising from correlation analysis of *Primary* and *Secondary* data, was considered appropriate for enquiries in 'education and the social sciences because it allows for the measurement of a number of variables and their relationships' (Cohen *et al.*, 2000, p.199). The peer assessments taking place in this study suggest a number of relationships that might reveal interesting findings if subjected to correlation analysis; for example, the relationship between the consistency of intra-peer agreement of a band and the inter-peer assessment of its performance; the self assessment of free-riders and their final grade. Correlation analysis involving peer as-

sessments that has been subjected to statistical tests may, however, be seen as being too detached from the natural setting and not 'grounded' on the raw data; stripped from the particular context in which individuals awarded their marks. Although such correlation analysis might indicate general relationships between variables, this could be at the expense of context, which is important for naturalistic enquiries such as this. Appropriate caution, therefore was taken, when examining the interpretation of the correlation analysis. As such, it provided a background role only, to more useful *Primary* and *Secondary* quantitative data and to the qualitative interview data, which was of prime importance to the study.

Correlation analysis of the *Primary* and *Secondary* data types was described as Tertiary data. As before, this is explained further in Chapter Five, which considers the analytical methods employed.

As indicated in Table 3.1 (p.103) with a Performance Management population size of 170 in the study and a sample size of 100%, the information generated from the marking data across each of the three data types was detailed and extensive. Deciding on an appropriate method to unify, analyse and interpret the intra-peer assessment data required careful consideration. Harnessing and utilising the potency of grounded theory for such purposes was considered a possible approach. This is discussed in the following section where it is explained how the principles of grounded theory analysis were adapted to the quantitative assessment data.

3.6.5.1 Other variables

This enquiry suggests a number of other variables, relating to the students involved in the research and from the peer assessment data itself that required careful consideration. Demographic information, describing a range of potential variables, appears in the Introduction and elsewhere, including age, gender, prior experience, and proficiency on particular instruments. Deciding whether or not to include such variables required a review of the context, aims and scope of the research. As such, it was important to consider the following points.

At the outset of the enquiry the overarching aim was an exploration of approaches to peer assessment that might support learning. The focus of the study involved the use

of students' personal attributes as intra-peer assessment criteria as an approach towards developing those qualities. It was also envisaged that the five-year action research period would culminate in the creation of a new model of peer assessment.

The data generated from the action research would, in all probability, be extensive and detailed as a consequence of the substantial length of time in the field. It was important for the aims and scope of the enquiry that resources should be concentrated on the area of personal attributes and intra-peer assessment in order to offer a study that had sufficient depth and meaning.

Investigating gender differences would be interesting when, for example, comparing students' personal attributes that were formulated for each sex, differences in peer marking and interview responses analyses by gender. However, this implied a substantial undertaking in which its potential merited a separate investigation. I was also concerned that there would be a gender imbalance, as courses in popular music typically attract low numbers of female students and might lead to a study that was not representative or meaningful.

I also decided not to incorporate age as a variable in the research analysis. From common experience, age is very rarely a factor for musicians performing in ensembles. Of much greater importance for a band is the ability and breadth of experience among its members.

To include instrumental expertise, or type of instrument performed by students, in the research analysis was disregarded on grounds of it being only marginally relevant to the context of the enquiry. The focus was that of developing personal attributes in rehearsing in a band, rather than developing instrumental proficiency.

Prior experience, although significant for musicians rehearsing in a band would, I considered, pose problems in defining the characteristics of such a variable. This might be determined, for example, on previous academic experience including vocational qualifications such as the BTEC National Diploma. Alternatively, it could be on the basis of substantial 'real world' experience in rehearsing and performing. As all students in the cohort were at the same starting point for their Performance Man-

agement band work and also their status as undergraduates, this indicated a more suitable chronology of their peer assessment experiences.

Finally, the literature about peer assessment in music that was surveyed in Chapter One did not reveal age or gender as key variables in the relevant studies. More apparent, however, was a significant interest in individuals and their particular characteristics that often become manifest in group work, especially the free rider and the shooting star. This became particularly noticeable when examining intra-peer assessment data. It was decided to incorporate free riders and shooting stars as key variables arising from the data in the research analysis. The grounded nature of the data and its analysis is also described on page 179.

3.6.6 Deciding on the method for analysing the quantitative assessment data

As discussed in Section 3.6.3, this study was not concerned with a particular theory or hypothesis in which data was specifically generated in order to confirm or disprove; it was characterised more by inductive thinking and reasoning. Data in this study was simply created as the lines of enquiry proceeded, from which data patterns, leading to propositions, might emerge. With emergent data in an enquiry of this type, I considered that using and adapting grounded theory to unify and analyse the quantitative assessment data wholly appropriate. Although grounded theory is typically employed in studies involving qualitative data, it is also considered equally suitable for quantitative contexts. For example, the first studies contained in the seminal text on grounded theory by Glaser & Strauss (1967) employ quantitative data while Strauss & Corbin (1998) make explicit their views that grounded theory is a general method applicable to both qualitative and quantitative investigations. Using grounded theory for practice-based research enquiries (as exemplified in this thesis) is, as Piantinida & Garman (1999) note, an established practice in the field of education.

Grounded theory is considered suitable for analysing extensive and detailed data, where its particulars might be lost through the use of powerful statistical tests alone. It is also a method that can be used to analyse and unify large quantities of information such as that generated through this research.
As the case for employing grounded theory became increasingly compelling, I decided to adapt its procedures to the analysis of the quantitative assessment data. The

adaptation was founded on its key principles: simplicity and fitness for purpose (Cohen *et al.,* 2000), appropriateness of the data context and maintaining the numbers and the words together in the ensuing analysis so 'one never strips the data at hand from the contexts in which they occur' (Miles & Huberman, 1984, p.21). The following paragraphs describe how I adapted grounded theory principles in this study.

As previously explained, the data was defined using three categories: *Primary*, *Secondary* and *Tertiary*. The term *Primary* was used to denote the intra-peer assessment marking data; *Secondary* was used to denote data that was subjected to low-powered statistical tests; *Tertiary* denoted data arising from correlation analysis involving *Primary* and *Secondary* data. This hierarchical structure was devised in order to indicate the relationship of each data type arising from the 'grounded' quantitative assessments. Thus, Primary data, comprising the raw intra-peer assessment marks that students awarded (and propositions resulting from the analysis of these), claimed the strongest status; Secondary data, resulting from low-powered statistical tests (and propositions deriving from these) was regarded as having less strength; Tertiary data (correlation analysis arising from the statistically tested Secondary data, and Primary data), had the weakest status.

The coding process for the quantitative assessments was built upon the principle of simplicity and that of striving to ensure that analysis remained 'grounded' on the data. Open, axial and selective coding procedures, following grounded theory traditions, were employed. Specific coding processes were developed also in order to further analyse the quantitative data. One of these processes employed simple three-way category descriptors: *high, moderate* or *low*. Another procedure was developed in order to categorise comparisons between the quantitative data; this was also coded through simple three-way descriptive categories: *greater than, equal to* or *less than.* Defining the measurement boundaries (for individual normalised difference coefficients, for example, *high* being defined as *0.90 or above*) was determined post-data rather than being specified in advance. Defining the measurement boundaries of the categories *after* being acquainted with the assessment data, followed grounded theory traditions that coding categories *arise* out the data themselves, rather than being prescribed in advance.

Open coding procedures were conducted on the primary and secondary data; *axial coding* produced subcategories derived from the primary, secondary, tertiary and

process features information (data associated with private or collaborative assessment; intra-peer assessment conducted before or after receiving performance mark; and inter-peer panel marks). *Selective coding* integrated the axial coding categories in order to provide a central focus of the assessment data. This integrative focus resulted in a number of *propositional claims* that were categorised by their strength, according to the status of the data as previously explained (i.e. *Primary*, *Secondary* or *Tertiary*).

3.6.7 Quantitative data collection: inter-peer assessment data

The performance marks that were awarded to the bands by the peer panels continued to be collected throughout the enquiry, as explained in Section 3.1, in order to compare and correlate peer panel marks with those of the tutors. An examination of this would help towards forming a judgement about whether the peer panels had developed sufficient skills and knowledge with which to award marks to bands performing at gigs. If it were established, from the correlation data, that the peer panels could assess bands and mark satisfactorily, then this would provide encouragement for further research involving peer assessment in this area of the curriculum.

Interrelationships arising from marks awarded by peer panels, tutors and the intra-peer assessments were also examined. Two non-parametric statistical tests were selected following the guidance offered by Siegel & Castellan (1988): Pearson Product Moment Coefficient of Correlation; Spearman Rank Correlation, A summary of this analysis was reported in Hunter (2004). It was found that there was a strong consistency of agreement in marking between the peer panels and that of the tutors. This line of enquiry was discontinued after phase one, because of the refocus on intra-peer assessment and process features as previously explained.

3.6.8 Qualitative and quantitative data collection: the process features enquiry

A discussion of the process features of peer assessment that were considered important can be found in Section 2.3.3. Interventions in the rehearsing and performing cy-

cles generated data relevant to the process features enquiry. This generated qualitative interview data and quantitative assessment data related to the following process features.

Timing of intra-peer assessments within the cycle

Interventions in the *timing* of intra-peer assessment were explored through a number of rehearsing and performing cycles. Bands conducted their intra-peer assessments at two different places within the cycle (either before their on stage performance or after their on stage performance) and their responses arising from these circumstances appear in the interviews with the 16 *Performance Management* students.

Private or collaborative intra-peer assessment

Conducting intra-peer assessment in either a private or collaborative setting was explored through various interventions (intra-peer assessments conducted privately, prior to bands receiving their performance marks; intra-peer assessments allowing collaboration, after bands receiving their performance mark). Students' responses to intra-peer assessment under each of these conditions were collected from the 16 individual interviews.

Formative and summative marking

Students' responses to their formative marking feedback and their summative marking feedback arising from cycle 12 were collected from the six individual interviewees who were involved.

Self assessments

Students were invited to award marks to themselves as well as to their band members, during intra-peer assessment (although individuals were aware that such marks would not count towards their summative assessment). Interviews with the *Performance Management* students provided an opportunity to discuss their self-assessments, especially in relation to the marks awarded by their bands.

3.7. Standards of quality and verification

Naturalistic enquiries have implications for establishing appropriate criteria against which to evaluate the quality of the research. Key texts in the literature on formulat-

ing appropriate criteria include those of Lincoln & Guba (1985), Miles & Huberman (1984, 1994) and the survey conducted by Creswell (1998). Lincoln & Guba suggest a criterion of 'trustworthiness' as an approach towards establishing standards of quality and verification. The concept of 'trustworthiness' is now firmly established in the literature on criteria as an approach that is appropriate towards ascertaining standards of quality and verification for enquiries characterised by naturalistic settings and constructivist thought (Lincoln & Guba, 1985, p.114; Creswell, 1998, p.200). 'Trustworthiness', following the account of Lincoln & Guba, comprises four criteria: credibility, transferability, dependability and confirmability. These are based on the following techniques: prolonged engagement, persistent observation, closeness to participants, peer debriefing, member checks, triangulation, providing 'thick descriptions', dependability and confirmability audit.

Credibility is a term used by Guba & Lincoln to describe verification through, for example, prolonged engagement, persistent observation, closeness to the participants in the study, peer debriefing, member checking and triangulation. Each of these was evident in this enquiry: the research in the field was conducted over a period of five continuous years with persistent observation throughout all nineteen rehearsing and performing cycles. Close participation was established also through my role as the module tutor to the students throughout this period of time. For example, it was calculated that my contact time with the *Performance Management* students during the period, was in excess of 250 hours. Peer debriefing comprised extensive and regular, often monthly, discussions with my two supervisors; some of their colleagues provided additional feedback and briefing of my analysis and written work. Topics arising from this research have been publicly presented at Higher Education forums, including conferences; findings have been published in scholarly, peer-reviewed publications. Four of the sixteen students who were interviewed were also involved in 'member checking' of the interview transcripts and coding. Although triangulation of data was not anticipated in the study, several findings, described later in Chapter Five, are also supported by triangulation of quantitative and qualitative data.

Transferability, Lincoln and Guba suggest, is a naturalistic parallel of external validity, but it is not for the researcher to provide an index of this. Researchers, they maintain, should provide evidence of transferability through rich data and 'thick descriptions' (Geertz, 1973). Both qualitative and quantitative data in this study can be re-

garded as fulfilling the criteria of being a rich and thick data description. The appendix also contains examples of the following: interview transcript; interview data analysis procedures; illustrations involving interventions in the cycles informed by interviews; quantitative assessment data analysis; open coding analysis; axial coding analysis; selective coding analysis. Transcripts of all sixteen interviews and over 200 pages of quantitative data description and analysis are maintained separately.

Dependability is demonstrated through the methodological techniques that were conducted including: whether they were applicable to the research; whether they were applied consistently; whether there existed an appropriate time scale and samples (Lincoln & Guba, 1985). The time scale of this research was sufficient to allow an appropriate action research methodology to emerge and develop that ensured congruence between the research enquiries and the methods that were used for the investigation of these. In terms of using appropriate samples, this criterion was exceeded in the research, consisting of the entire population of each Performance Management cohort. The interview samples comprised participants who had appropriate experience of intra-peer assessment; for the 2002/3, 2003/4 and 2004/5 cohorts the sample sizes were 24%, 26% and 42% respectively.

Confirmability was demonstrated through a confirmability audit comprising the entire period of the research (2000-2007). It consists of the notes to all of the tutorials with my supervisory team including saved email correspondence, samples of grounded theory coding and analysis checked by my supervisory team, complete audio recordings of all interviews, complete video recordings of all Performance Management performances, all my rough notes, and every item of work that was submitted to and returned by my supervisors throughout the research.

In the accumulation of these, it is hoped that the criteria for establishing the trustworthiness of this research has been satisfied. Lincoln & Guba however believe that 'criteria of trustworthiness are open-ended' and that 'no amount of member checking, triangulation, persistent observation, auditing, or whatever can ever compel; it can at best persuade' (p. 329).

Summary

This naturalistic study of undergraduate popular music group work and peer assessment is informed by constructivist theory. The methodology involved an investigation of three lines of enquiry: *intra-peer assessment; inter-peer assessment; process features of assessment*. The students comprised year one undergraduates rehearsing and performing in their bands and intra-peer assessing each other. Year three students participated in peer panels in order to assess the first year band performances. An action research approach was developed and several research interventions were conducted during the rehearsing and performing cycles. The enquiry later refocused itself on two remaining lines of enquiry: *intra-peer assessment* and *process features of peer assessment.* Peer assessment training activities were provided together with establishing basic principles for those involved, including its rationale, transparency and agreement. Students' personal attributes, in the context of group rehearsing and performing, were formulated and employed as intra-peer assessment criteria. Quantitative assessment data and qualitative interview data were gathered. Grounded theory was adapted to analyse each of these two data types. The chapter concluded with an explanation of the standards of quality and verification that applied to this research, following Lincoln and Guba's account of 'trustworthiness'.

Chapter Four: The Rehearsing and Performing Cycles

4.1 Illustrations of the cycles

4.2 Rehearsing and Performing Cycles: 2000-2001

4.3 Rehearsing and Performing Cycles: 2001-2002

4.4 Rehearsing and Performing Cycles: 2002-2003

4.5 Rehearsing and Performing Cycles: 2003-2004

4.6 Rehearsing and Performing Cycles: 2004-2005

4.7 Summary and analytical overview

4: The Rehearsing and Performing Cycles

This chapter explains the development of the rehearsing and performing cycles in the enquiry and analyses the action research interventions. These interventions enabled a variety of approaches to be explored and refinements to be made in the peer assessment activities. It produced a theoretical and conceptual framework for using personal attributes as assessment criteria, illustrated later in the chapter, in Table 4.1 (p.156) and Table 4.2 (p.157). A key element of this framework, which became clarified during the cycles, was the concept of employing *personal weaknesses* as intra-peer assessment criteria with which to assess each individual band member. The development of a theoretical basis for intra-peer assessment that supported learners during their rehearsing also became clearer. As the cycles evolved, it became apparent that supporting the individual necessitated a gradual change in the focus of the intra-peer assessment away from using group attributes[1] to that of individual attributes. A similar change involved a shift from using *self-determined* personal weakness attributes to *band-determined* personal weakness attributes for each individual. The use of personal weakness attributes, determined by the student's band rather than by him/herself, was a key underpinning of individuals' learning through intra-peer assessment activities during the cycles.

The purpose of this chapter, therefore, is to discuss the rehearsing and performing cycles employed in this study and to explain the action research 'interventions' (i.e. changes to the learning, teaching and peer assessment methods) that were conducted within each cycle. Indeed, the research can be regarded as being particularly distinctive in its development, practice and management of the rehearsing and performance cycles. As explained in Sections 2.4 and 3.2, these cycles were the means through which the peer assessment investigation progressed, from cycle to cycle and from year to year. Each cycle presented an opportunity for an action research intervention in order to explore changes in particular aspects of peer assessment. The rehearsing and performing cycles, fundamental not only to the chronology of the enquiry, were also the means through which data was collected and, ultimately, in the acquisition and creation of knowledge about peer assessment and learning.

The interventions that were conducted in each cycle were informed by experience and

[1] Group-agreed attributes, as explained in Section 3.3.1.2.

knowledge acquired from previous cycles. For example, the first cycle of the 2002-2003 *Performance Management* academic year was cycle 11 (the *Christmas Party* assignment). This cycle, together with its interventions, were built on the experiences of the equivalent 2001 *Christmas Party* cycle, as well as that acquired from the other cycles of the academic year.

4.1 Illustrations of the cycles

In order to help explain the cycles, a series of illustrations (Figure 4.2 *et seq.*) are used to depict their key features and associated interventions. At the head of each diagram appears the particular research line of enquiry (i.e. intra-peer assessment, inter-peer assessment, process features). Summarised in boxes beneath each are the activities and interventions that took place in each cycle. A dotted lined box indicates that its contents consist of an intervention or change from the preceding cycle. Finally, a series of arrows indicates the chronological development of the lines of enquiry taking place during the cycles. The conclusion of each academic year was a natural place in which to review, reconsider and refine aspects of the cycles and the action research. At this point, it may be helpful to revisit Figure 2.2 (p.88), which illustrated the action research framework. This framework, shown as Figure 4.1 (p.137), also indicates (in **bold**) the extent to which the cycles occupied a central position in the enquiry.

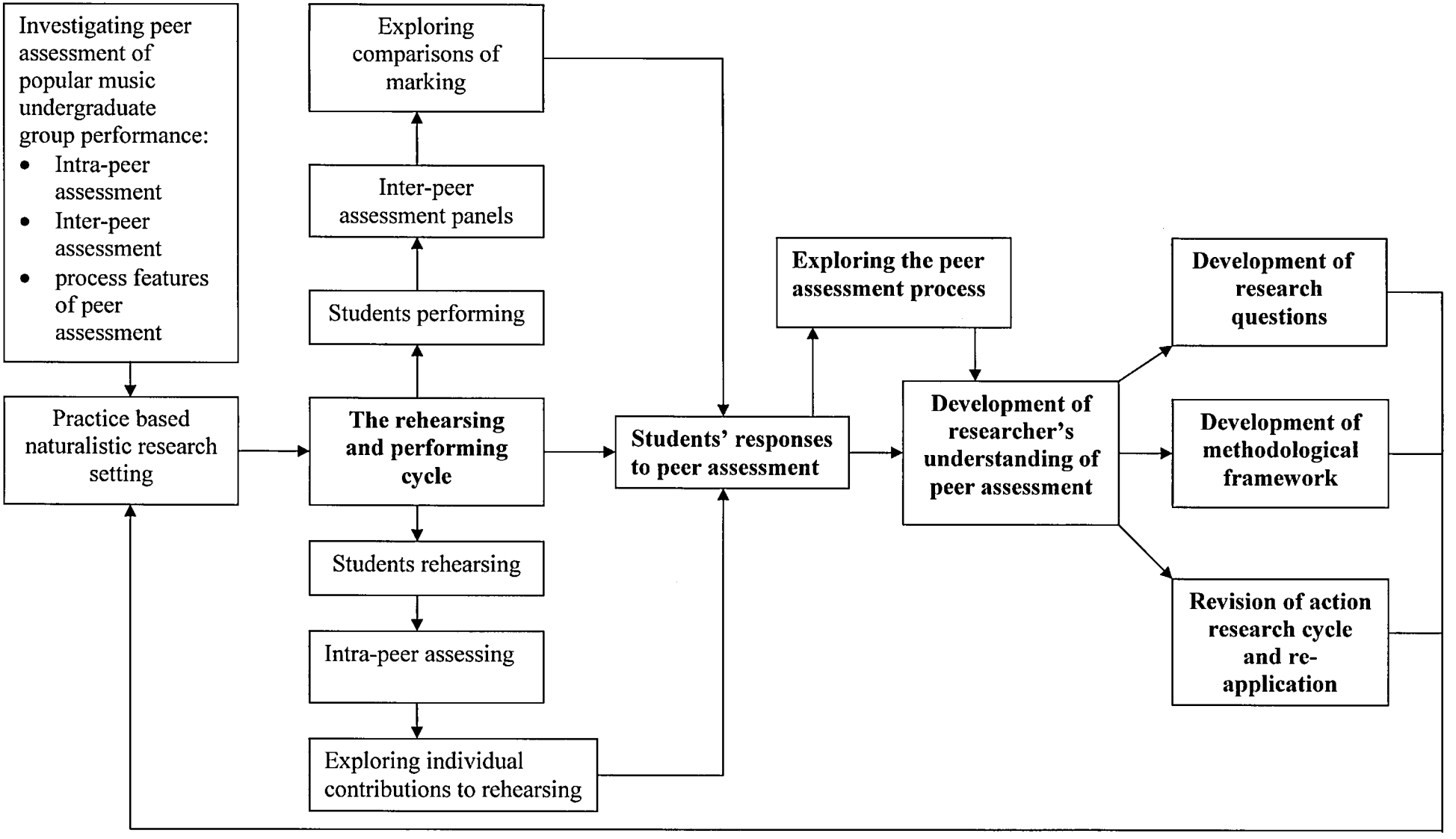


Figure 4.1 The rehearsing and performing cycles in the action research

4.2 Rehearsing and Performing Cycles: 2000-2001

As explained in Chapter Three *Methods of Enquiry,* the research was characterised by three lines of enquiries: intra-peer assessment; inter-peer assessment; process features of peer assessment. Figure 4.2 (p.140) illustrates each of these together with the related peer assessment activities and the date of each cycle.

For the intra-peer assessment enquiry, a general holistic criterion *contribution to rehearsals* was employed for each cycle, with marks calculated using the 'zero-sum' method, as explained in Section 1.4.4. No interventions were conducted for the intra-peer assessment line of enquiry during the year. Deciding against interventions in this line of enquiry was for reasons of providing some stability in this first year of action research, in which acquiring tutor experience and confidence was the overriding aim. To initiate interventions across all three lines of enquiry at the outset of the research into peer assessment was, I thought, unwise.

For the inter-peer assessment enquiry, the bands themselves generated the assessment criteria for each of the three performances. In order to compare the panel marks with those of the tutors, a single peer panel assessed all the performances that were given in the first cycle. Following an encouraging response by the panel to marking performances and providing written feedback, two peer panels were formed in order to provide additional participatory opportunities for peer assessment of the next two performances. Each panel assessed different bands on consecutive evenings and the marks that they awarded were compared with those of the tutors.

The process features enquiry explored two aspects: the setting of the intra-peer assessment activity (awarding marks either in private or through a collaborative discussion), and the timing of intra-peer assessment in the cycle. The first two intra-peer assessments were conducted in private and took place after students were aware of their band's performance mark. For the next assessment however, this was changed whereby students were given the option of collaborative assessment with their band in order to agree the marks that they awarded to themselves. The timing of the intra-peer assessment activity was also changed; being conducted before bands had knowledge of their performance mark.

There was a total of six rehearsing and performing cycles during the period, comprising three BA1 Performance Management and three Performance Studies assignments.

Summary of key points and analysis informing the research:

- Obtaining experience and confidence in using peer assessment was the overriding research objective in this first year.
- No interventions were conducted for both intra- and inter-peer lines of enquiry because it was considered important to establish and maintain stability during the introduction of the peer assessment system.

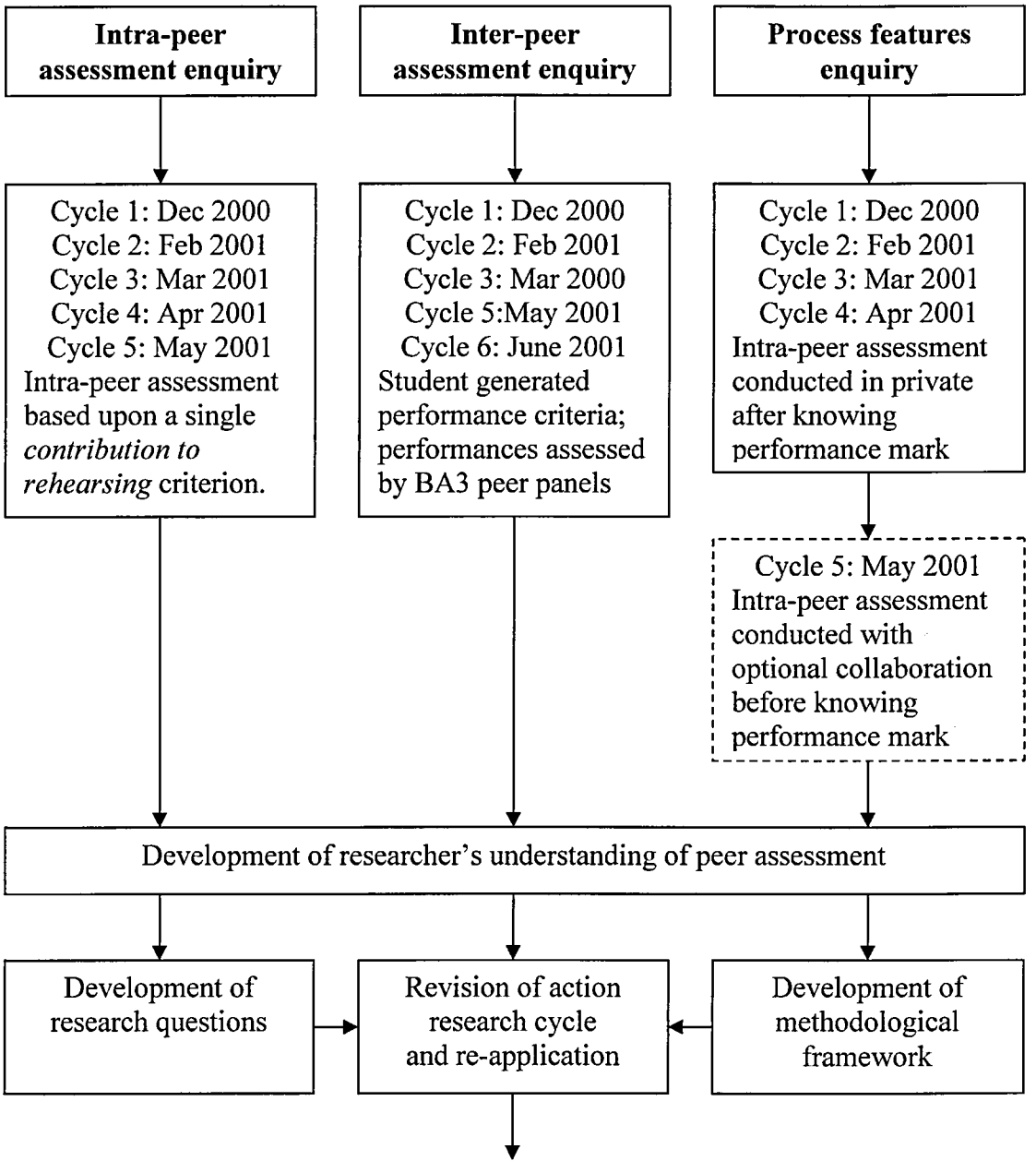


Figure 4.2 Development of the rehearsing and performing cycles: 2000-2001

4.3 Rehearsing and Performing Cycles: 2001-2002

Reflecting on the experiences of this initial year, in which useful experience of peer assessment was being acquired, I decided to begin exploring ideas about peer assessment that might support particular aspects of learning. The idea of using students' personal attributes as assessment criteria began to take shape during this second year. The purpose of employing personal attributes as peer assessment criteria was that of developing students' self-awareness and an awareness of others' rehearsal attributes.

Figure 4.3 (p.143) illustrates the development of the cycles during the 2001-2002 academic year. For cycle 7 (the 2001 *Christmas Party* assignment), as explained in Section 3.3.1, each student formulated three personal attributes that they considered were *important to their rehearsing* which they agreed could be used as assessment criteria. Additionally, each band established three group attributes (i.e. *group agreed attributes*) arising also from what they considered *important to their rehearsing.* A total of six personal attributes for each student, therefore, were used as intra-peer assessment criteria for their rehearsing.

The same six personal attributes for each individual, was used again as assessment criteria for the next cycle 8 (the *Venues & Audiences* assignment). In view of the literature on category-weighted criteria (surveyed in Sections 1.4.2.2 and 1.4.2.3), I decided also to explore the use of criteria that attracted an equal assessment weighting. For cycles 9 and 10, therefore, each personal attribute criterion was weighted at one sixth of the total. Another similar intervention in this cycle involved inviting Performance Management and Performance Studies students to produce category-weighted band performance criteria for the peer panels. The purpose of this was to examine how peer panels might respond to using category-weighted band performance criteria, compared with their previous experience of using non-weighted criteria.

There were also interventions in the inter-peer assessment line of enquiry. In order to compare the marks awarded by the panels with those of tutors in cycle 7 (*Christmas Party*) and cycle 8 (*Venues & Audiences*), peer panels assessed five of the ten bands who performed on the first *Christmas Party* evening gig, while tutors assessed the remaining five bands who performed on the following evening. This arrangement was reversed for the subsequent *Venues & Audiences* performances: bands that were

previously assessed by a peer panel became tutor assessed, and vice versa. As the bands' membership remained the same for both assignments, it was thought that from this interchange of assessors, interesting findings might arise from the correlation analysis. For the final cycle, a further intervention was made involving the process through which peer panels awarded their marks. On previous occasions, panels discussed the performances and decided together, in a collaborative setting, the marks that should be awarded. However, in order to investigate the consistency of marking agreement within peer panels, each panel member agreed to submit their marks without discussion. The average of these marks was awarded to the band.

For the process features enquiry, intra-peer assessment activity reverted, for cycle 7, to being conducted in private. This was changed to optional collaboration for cycle 8, in order for each setting to be compared, as I considered it useful for students to experience marking conducted in each of these situations. The intra-peer assessment activity, for both cycles, was conducted before bands were aware of their performance mark.

There was a total of four rehearsing and performing cycles during the period, comprising three *Performance Management* and one *Performance Studies* assignment.

Summary of key points and analysis informing the research:

- The use of students' personal attributes as intra-peer assessment criteria was introduced.
- Students experienced conducting the intra-peer assessment activity in both private and collaborative settings, and also at different times in the cycle.

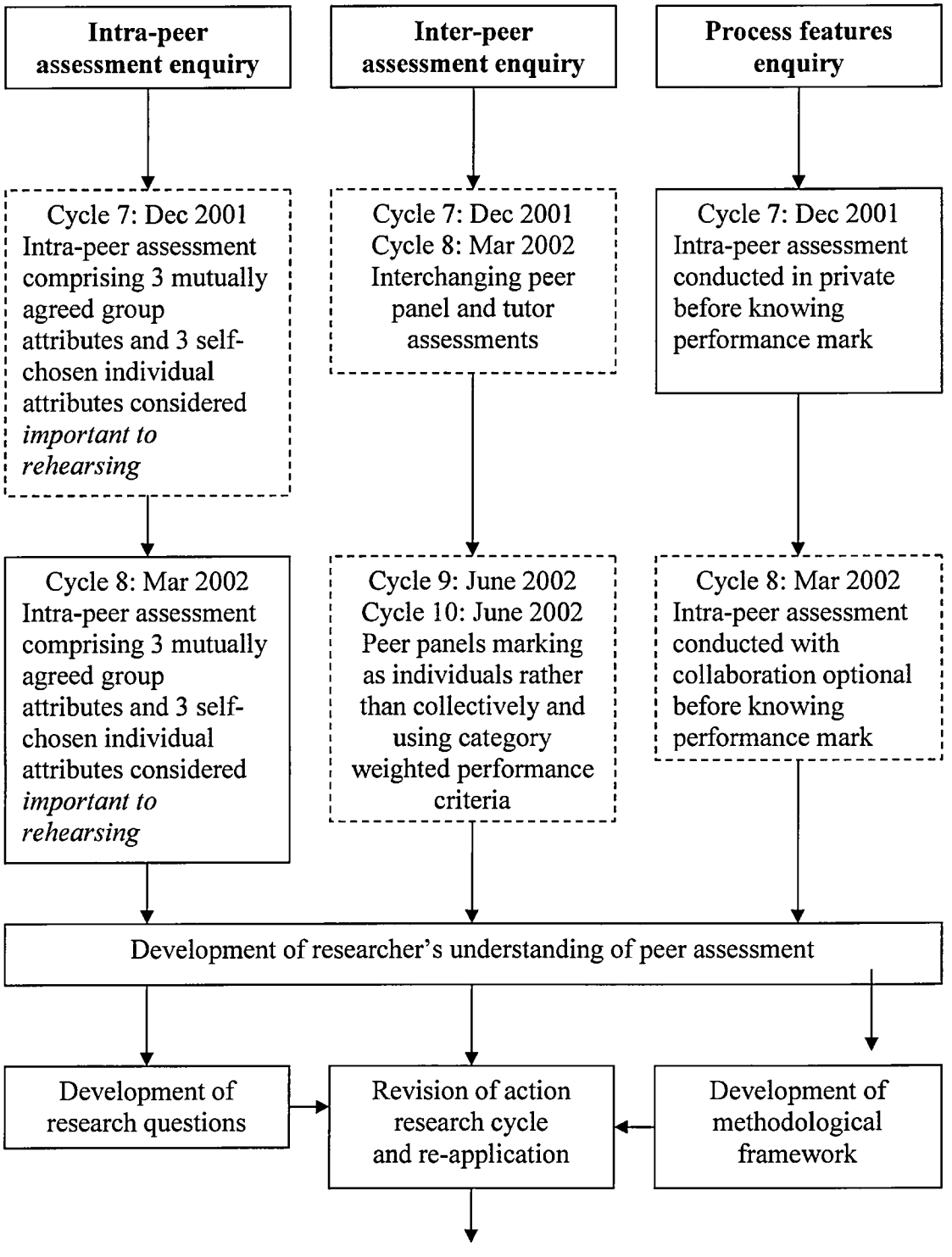


Figure 4.3 Development of the rehearsing and performing cycles: 2001-2002

4.4 Rehearsing and Performing Cycles: 2002-2003

It was explained in Section 4.4, that the enquiry was refocused for the 2002-2003 academic year, in order to concentrate on particular aspects of peer assessment. As a consequence of this, the inter-peer assessment enquiry was discontinued and the *Performance Studies* bands were no longer involved in the study. Figure 4.4 (p.147) illustrates the rehearsing and performing cycles during this period with the two remaining lines of enquiry (i.e. intra-peer assessment and process features of peer assessment).

There were important interventions involving the intra-peer assessment enquiry. To promote self-awareness of their personal attributes in rehearsing, students were asked to reflect on their *personal strengths* and *personal weaknesses* in rehearsal. For cycle 11 (*Christmas Party*), students provided two personal strengths attributes and one personal weakness attribute that they agreed could be used as intra-peer assessment criteria. The three group-agreed attributes continued as before. It was thought that inviting each student to offer just one personal weakness within the total of six attributes would help to avoid an undue emphasis on negative qualities. Given that the formulation of these attributes took place at an early stage in their undergraduate studies, there were risks involved in expecting individuals to be able to freely disclose *personal weaknesses* to their peers, of whom many would be unfamiliar. Moreover, would students be in agreement to their *personal weaknesses* being used as intra-peer assessment criteria for their peers when disclosing these?

Students' responses to intra-peer assessment using personal strength and weakness criteria in cycle 11 were generally positive and this encouraged further interventions in this area. As I acquired confidence in facilitating the use of students' *personal weakness* attributes, I decided to focus their awareness of these a little further. Developing an awareness of their *personal weakness* attributes, together with those of their band members' would, I considered, provide important learning experiences. Accordingly, for cycle 12 (*Venues & Audiences),* the number of students' *personal weakness* attributes was expanded from one to three. Students remained in the same bands for both of these assignments and it was agreed to re-use these attributes that they had formulated for the final cycle of the year: cycle 13 (the *Decades* assignment). Repeating these six attributes (three *personal weaknesses* and three group-agreed attributes) together with maintaining

the same bands from assessment to assessment would also enable interesting comparisons to be made, both in the quantitative assessment data, as well as that arising from interviews.

Although by this stage, peer panels were no longer a part of the research enquiry, these panels had, nevertheless, become a normal and popular feature of band assessment brought about by the action research. Students, from the comments they made in the group peer panel interview, valued their involvement in peer panels, particularly in the formulation of their feedback. Indeed, the year three *Performance Techniques* students were occasionally entrusted to assemble panels entirely by themselves.

A number of interventions took place also in the process features enquiry. Bands conducted their intra-peer assessment for the *Christmas Party* and *Venues & Audiences* cycles after being aware of their performance mark. The timing of this was changed, however, for the final *Decades* cycle and the assessment was conducted before bands were aware of their performance mark.

As in the previous cohort, there were interventions involving the intra-peer assessment setting in order to enable students to experience both private and collaborative marking activities: private marking for the *Christmas Party* and *Venues & Audiences* cycle; optional collaborative marking for the *Decades* cycle.

A final intervention involved providing a formative intra-peer assessment midway through the *Venues & Audiences* cycle. The purpose of this was twofold: firstly, it explored whether the summative marks would be higher than their formative marks; secondly, it enabled an evaluation of the practicability of providing formative feedback at the midpoint of the cycle, given that students would have only three more weekly rehearsal sessions in which to respond to this.

There was a total of three rehearsing and performing cycles during the period, comprising three Performance Management assignments.

Summary of key points and analysis informing the research:

- Refinements were made in the use of personal attributes as assessment criteria, which focused on students' self-selected personal strengths and weaknesses, in order to develop their self-awareness in these areas.
- Formulating intra-peer assessment criteria based on students' personal weaknesses posed a risk of damaging the peer-to-peer and peer-to-tutor relationship.
- There was a positive response, in general, to the use of personal weakness attributes; self-selected personal weakness criteria were increased, with students' agreement, from one to three.
- Students continued to experience intra-peer assessment in different settings, and at different times in the cycle.

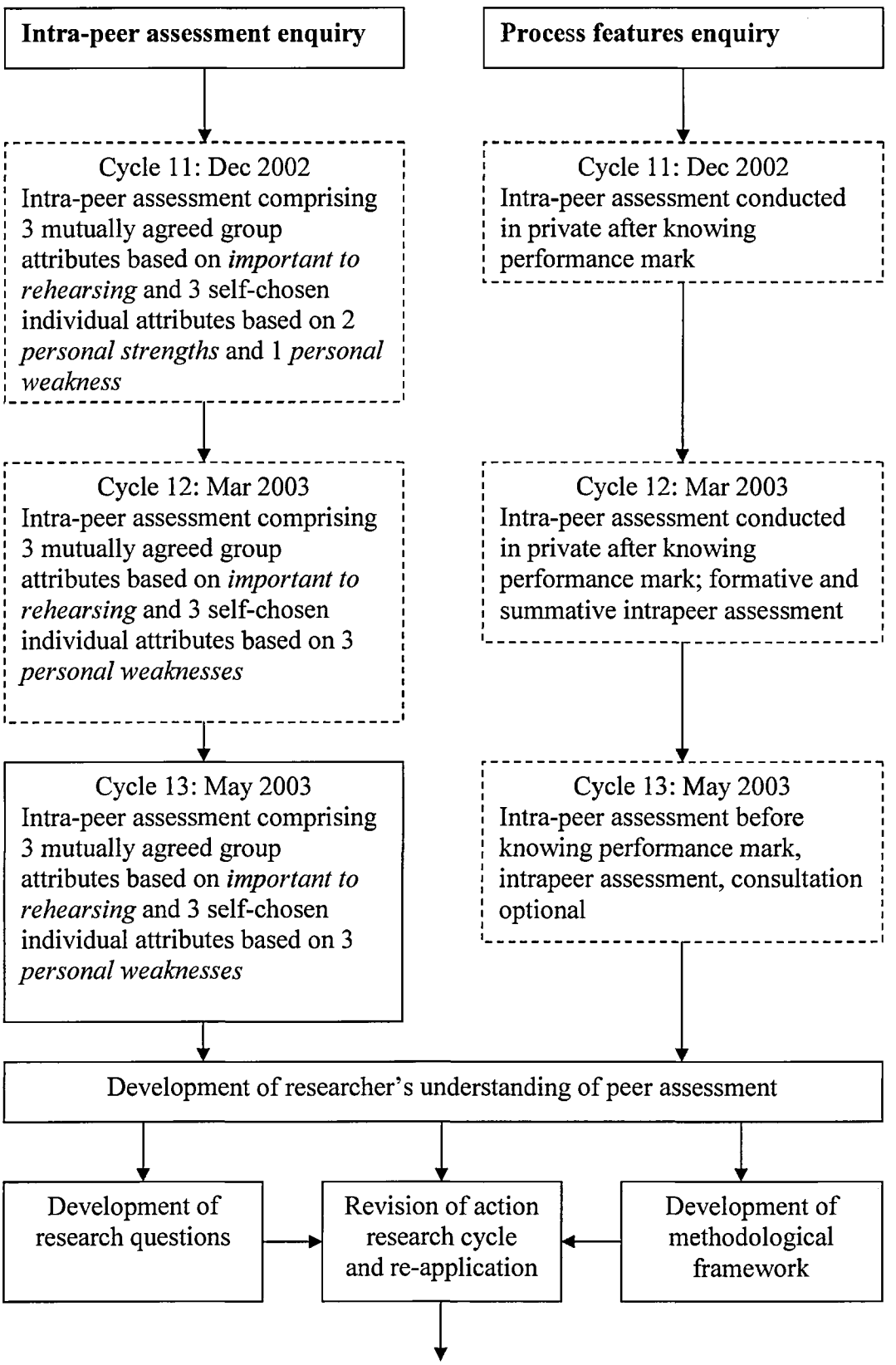


Figure 4.4 Development of the rehearsing and performing cycles: 2002-2003

4.5 Rehearsing and Performing Cycles: 2003-2004

Exploring the use of students' personal attributes as intra-peer assessment criteria for rehearsing continued during the 2003-2004 cohort and these involved significant interventions within the *Christmas Party* assignment of cycle 14. Figure 4.5 (p.151) illustrates the rehearsing and performing cycles during this period. Whereas in the previous year, the establishment of three *personal weakness* attributes was introduced in the *Venues & Audiences* assignment, I was sufficiently confident that the 2003-2004 cohort was ready to engage with this activity for their *Christmas Party* assignment.

As before, I considered using students' *personal weakness* attributes for assessment criteria as a risk worth taking. Because the experience from the previous year was generally positive, this factor, together with my increasing confidence in using peer assessment, contributed to my decision to introduce personal weakness attributes criteria early in the academic year. The underlying intention was that of engendering students' self-awareness of these areas as soon as I considered it possible. Deciding when it was possible appeared to involve a dependency on factors such as the readiness or maturity of students arising from the discussion with them about the use of their personal weakness attributes as intra-peer assessment criteria. My knowledge of the class and of individuals within it was important towards forming a decision about *when* to introduce personal weakness attributes. The 2003-2004 cohort of students appeared to exhibit a readiness and maturity early during the Autumn Term especially in their embracement of the potential for wider feedback that I found during their band rehearsals.

For cycle 15 *Venues & Audiences*, I decided to explore an idea that emerged from interviews with Performance Management students. This involved inviting each band to suggest, between them, the three personal weaknesses that they considered were most appropriate for each individual band member. As students had already given considerable time rehearsing in their bands during the previous cycles, these band members would have acquired a good knowledge of each other's rehearsal qualities. I considered it therefore not unreasonable that the band with which students had previously rehearsed, should formulate the personal weakness attributes for each member. It was important that the process was regarded as being fair and transparent and that each student was agreeable to the attributes that had been suggested for them.

In order to support transparency and the principle of student agreement, I decided that individuals who disagreed with any one or more of the personal weakness attributes chosen by their band were entitled to a veto. Where this occurred, the next step involved a discussion with the band in order to agree to suitable alternatives. Failing that, the student would have the right to re-using the three personal weakness attributes that they themselves formulated for the *Christmas Party* cycle.

Again, there were substantial risks involved in formulating personal weakness attributes in this way. At jeopardy, for instance, was the friendship between students, as well as the student-tutor relationship. During the process and in the period of time following it, I noticed no evidence of any such breakdown. Students appeared to respond positively to their personal weaknesses being decided by this method. There were some signs of polite embarrassment when, for example, one or two bands were finding it difficult to identify three personal weaknesses for individuals who they clearly regarded as *shooting stars*. Several bands were cautious in devising appropriate phrases to describe a personal weakness that avoided causing offence to the individual concerned. The session required careful tutor facilitation. It was suggested that individuals might prefer to exclude themselves temporarily from their band when it became the turn for their personal weaknesses to be determined. Where an individual band member decided to be present for the discussion, it was on the understanding that they had to remain silent in order to allow their band colleagues to reach their decisions without interference. In fact, several individuals decided to leave the room when it became their turn to be discussed. Perhaps this resulted from a combination of nervousness, discomfort, or that it would display trust in their band to decide, fairly, what was appropriate. I chose not to enquire about their reasons for deciding to leave as I felt that this would be an unwelcome intrusion, which might have added to any discomfort already being experienced. All band members reformed in order to agree on the three group attributes, as in previous cycles. To summarise, cycle 15 comprised six personal attributes: three *personal weaknesses* determined by the band for each student, and three mutually agreed *group attributes*.

For the following *Decades* cycle 16, I repeated the idea of inviting bands to decide members' *personal weaknesses*. I also decided to increase the focus on personal weakness attributes by reducing the three group attributes to one single tutor-imposed group attribute criterion: *general contribution to rehearsing.* Consequently, the previous

six criteria that were equally weighted at one sixth became four criteria equally weighted at one quarter, with the three personal weakness attributes accounting for three-quarters of the intra-peer assessment criteria. The purpose of increasing the weighting of personal weaknesses was that of providing further motivation for developing or improving these attributes.

As in the previous year, there were interventions in the assessment setting in order to enable students to experience both private and collaborative marking: private marking for the *Christmas Party* and *Venues & Audiences* cycles, optional collaborative marking for the *Decades* cycle. All intra-peer assessments during this year were conducted before students were aware of their band performance mark. Conducting the intra-peer assessment activity between the final rehearsal and the summative performance was by now, an established procedure and remained so until the end of the enquiry.

There was a total of three rehearsing and performing cycles during the period, comprising three *Performance Management* assignments.

Summary of key points and analysis informing the research:

- Self-selected *personal weakness* attributes were introduced earlier than in the previous year, in order to more speedily develop students' self-awareness of their personal attributes.
- Deciding when to introduce particular personal attribute criteria appeared to be dependent on factors such as students' readiness or maturity.
- Band-determined *personal weakness* attributes were introduced as assessment criteria for each individual although this posed risks to interpersonal relationships.
- Reducing the three group attributes to a single *general contribution to rehearsing* criterion increased the focus on personal weakness attributes.
- Transparency and seeking student agreement to the process, including the proviso of a veto, were key principles that emerged during these cycles.
- There was a relationship between increasing confidence of peer assessment and increasing risk-taking.

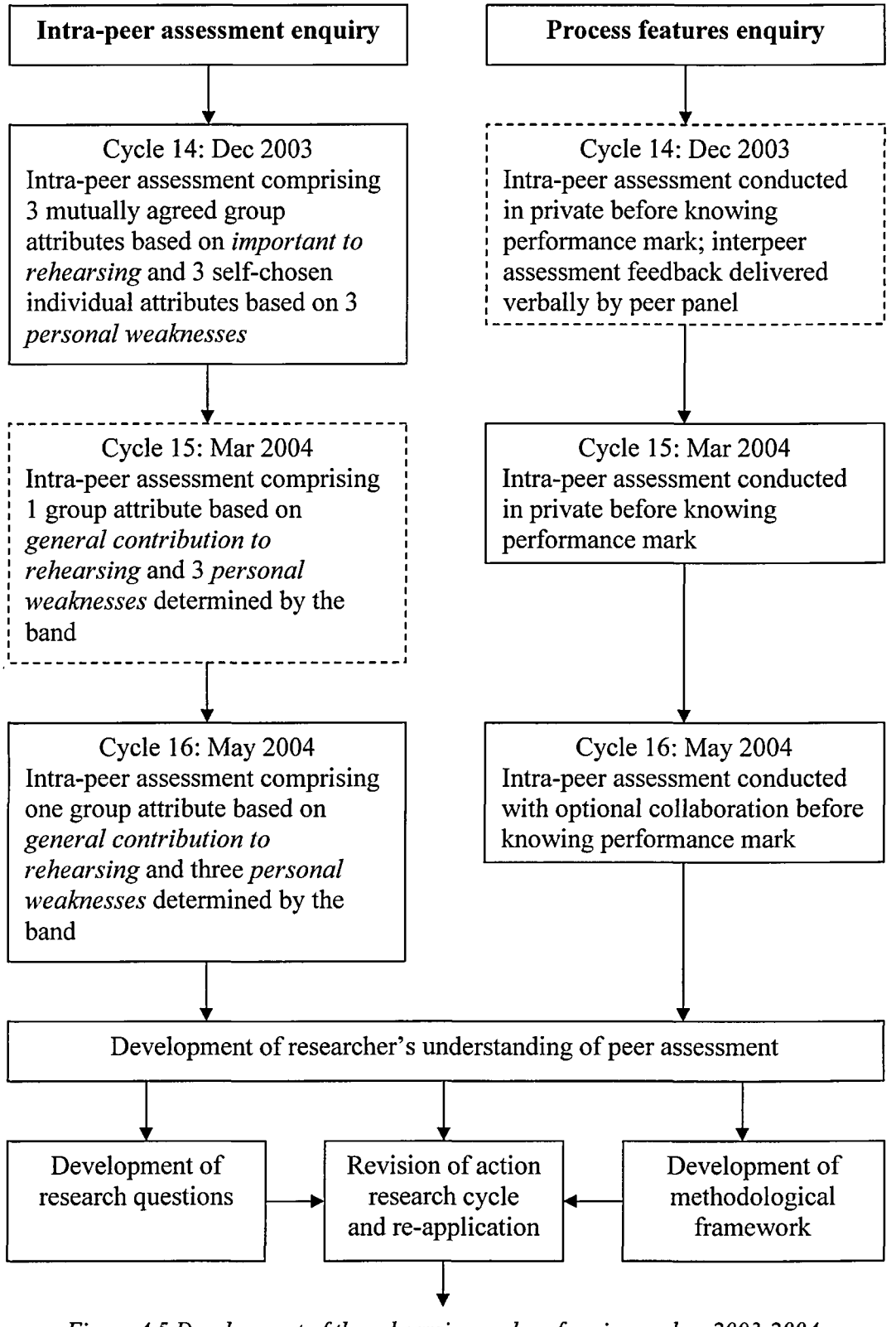


Figure 4.5 Development of the rehearsing and performing cycles: 2003-2004

4.6 Rehearsing and Performing Cycles: 2004-2005

This was the final academic year in which action research involving the rehearsing cycles was carried out; it is summarised in Figure 4.6 (p.154). In cycle 17 (the *Christmas Party)*, students selected three individual personal weaknesses, as in the equivalent 2003 *Christmas Party* cycle. For this present cycle, however, I decided to increase the weighting of students' personal weaknesses even further, by dispensing with the *general contribution to rehearsing* group attribute criterion that was previously used. As a consequence of this, intra-peer assessments were based entirely on students' three personal weaknesses, instead of the three-quarter weighting of the previous cycle. The readiness and maturity of the students were thought to be determining factors when considering the risks involved in this approach. This was a small cohort, however, which enabled greater tutor knowledge of the students than usual to be acquired. It was on the basis of tutor knowledge that the concentration on *personal* weakness attributes was implemented. In so doing, it was hoped that individuals might become further focused on their personal weaknesses in order to provide a motivational spur for their improvement of these.

For cycle 18 (*Venues & Audiences)* and cycle 19 (*Decades)*, as in the previous year, the bands with which each student had rehearsed, determined students' individual personal weaknesses (subject to the agreement of each student as before). The purpose of employing the same band determined personal weaknesses over these two cycles was that of enabling a longer period of time through which their personal weakness attributes could be assessed on two separate occasions. This allowed a comparison of band determined personal weakness assessments, which might raise interesting comments from students arising from their interviews.

Similarly, as in the previous year, there were interventions in the intra-peer assessment setting in order to enable students to experience both private and collaborative marking: private marking for the *Christmas Party* and *Venues & Audiences* cycles, optional collaborative marking for the *Decades* cycle.

For the *Christmas Party* and the *Venues & Audiences* assignments, students conducted their intra-peer assessments without knowing their band performance mark. This was changed in the final cycle whereby students' intra-peer assessments were conducted

after they received their band performance marks. Regarding this final year of action research, rather than initiating further significant interventions, I decided that a consolidation of the process features of peer assessment was more appropriate. There were a total of three rehearsing and performing cycles during the period, comprising three Performance Management assignments.

Summary of key points and analysis informing the research:

- There was an increased focus on students' personal weaknesses attributes through dispensing with the single *general contribution to rehearsing* group attribute criterion, in order to provide further motivation for students to improve those areas.
- The readiness and maturity of the students was increasingly thought to be determining factors when considering the risks involved in formulating personal weakness attributes as intra-peer assessment criteria.
- Students continued to experience the intra-peer assessment activity in different settings and at different times in the cycle: these informed the process features enquiry.
- An understanding of process features of peer assessment was consolidated further during this final year.

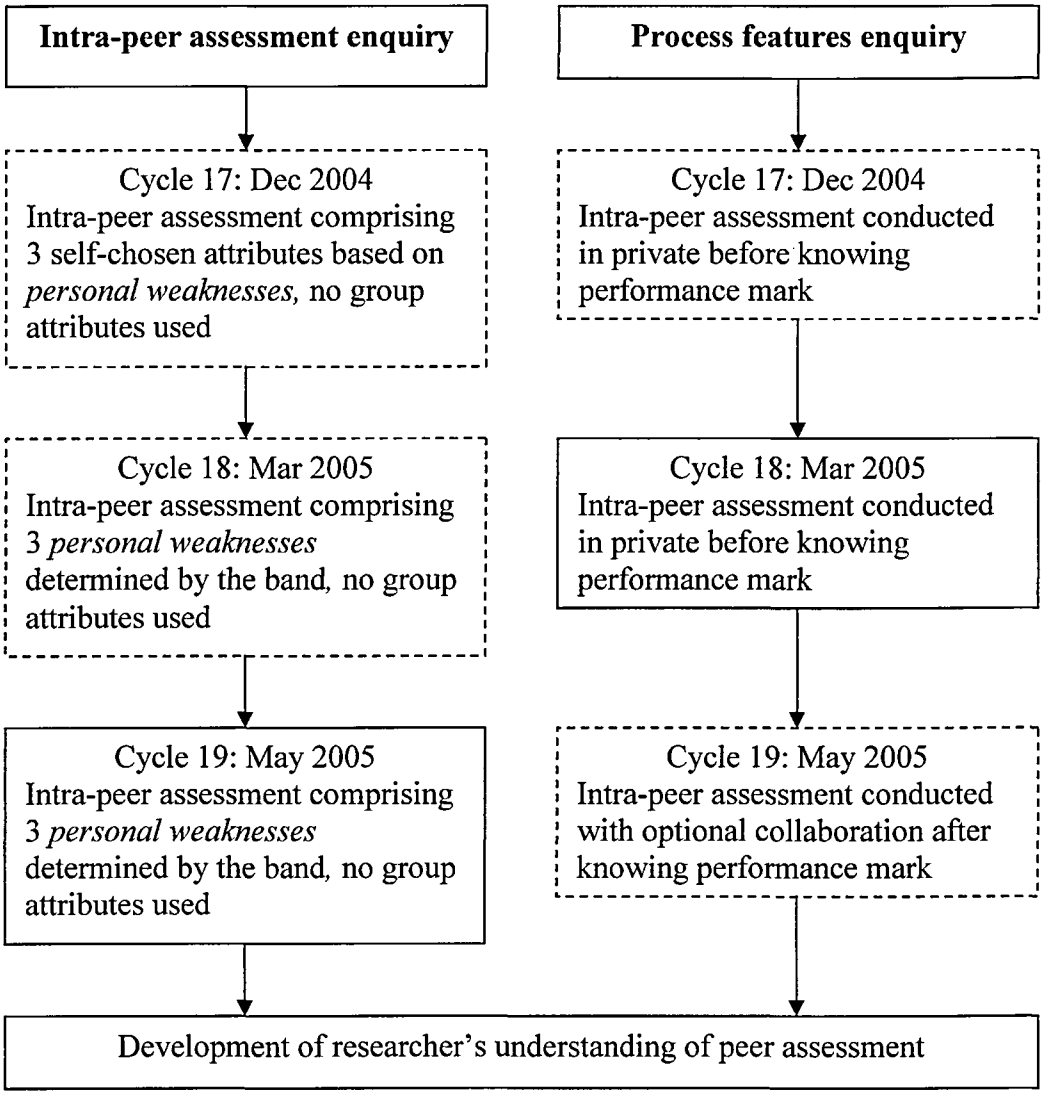


Figure 4.6 Development of the rehearsing and performing cycles: 2004-2005

4.7 Summary and analytical overview

This chapter described the nineteen rehearsing and performing cycles and explained the various developments that took place during 2000-2005. Illustrative diagrams were used in order to summarise the peer assessment activities of each rehearsing and performing cycle during the five-year period. The diagrams presented the three lines of enquiry that were investigated in the action research (intra-peer, inter-peer and process features of peer assessment). Appearing beneath each line of enquiry were the cycles and changes ('interventions') that were made to these, indicated by boxes surrounded by dotted line. The illustrations also made clear how these lines of enquiries became reduced from three (in cycle 1) to two (in cycle 11), and the consequential reduction of the modules involved in the enquiry.

The intra-peer assessment enquiry was primarily concerned with developing types of personal attributes to be used as assessment criteria. Exploring methods for establishing personal attribute criteria was a key activity in many of the cycles: individuals choosing these for themselves: bands deciding mutually agreeable group attributes: bands determining individual band member's attributes. A key approach was developed which involved the formulation of students' personal attributes based on self-identification of their *personal weaknesses*. Various combinations, numbers and permutations of personal attribute criteria, especially those based on *personal weaknesses*, were explored during the cycles. Many of the cycles employed six equally weighted personal attribute criteria.

The inter-peer assessment enquiry was primarily concerned with exploring the use of peer panels in order to assess band performances. Various data were collected during the cycles in order to compare peer panel marks with those of tutors; this is examined later, in Chapter Five.

A number of process features of peer assessment were also investigated during the cycles. The cycles were used to explore intra-peer assessment conducted in private and collaborative settings. Similarly, the timing of the intra-peer assessment activity was altered during many cycles in being placed either before or after the band performance.

The extent to which peer assessment evolved over the course of the cycles as a consequence of the action research is especially significant. This can be illustrated in

comparing the first cycle (intra-peer assessment simply based on a single holistic criterion *contribution to rehearsing),* with cycle 11 (three group attributes based on *important to rehearsing* and three self chosen individual attributes two *personal strengths* and one *personal weakness)* and also with cycle 19 (three *personal weaknesses* determined by the band with, notably, no group attributes being used).

The centrality of action research cycles to the enquiry and the refinements in peer assessment that took place as a consequence of the interventions is illustrated in Figure 4.1 (p.137). These refinements resulted in an increasing focus across a number of elements arising from the investigation. Table 4.1, presents these in the form of a model in which the types of personal attribute that were employed in the rehearsing and performing cycles are related to the focus on: the individual; risk; support for learning; motivation; readiness and maturity. This model, deriving from the rehearsing and performing cycles and the action research interventions, is considered further in Chapters Six and Seven.

Table 4.1 Model showing the relationship between the types of attributes employed in the rehearsing and performing cycles and the increasing action research focus

Types of personal attributes used as assessment criteria in the rehearsing and performing cycles	Increasing focus on the individual ▼	Increasing risk level ▼	Increasing potential for supporting learning ▼	Increasing motivational focus ▼	Increasing focus on maturity and readiness ▼	Increasing tutor confidence ▼
Tutor-imposed criterion: *'general contribution to rehearsing'*						
Group agreed attributes: *'important in rehearsing'*						
Individual attributes: *'important in rehearsing'*						
Individual attributes: *'personal strengths'*						
Individual attributes: *'personal weaknesses'*						

The model shows the evolution of personal attributes, which may be used as criteria for peer assessment, being established initially on a general or group basis, to that of being increasingly focused on individuals within a band. The first vertical column represents this *Increasing focus on the individual.* The next column indicates the increasing focus also in terms of the potential risks to interpersonal relationships that are posed by using particular types of attributes. For example, *general contribution to rehearsing* or *group attributes considered important in rehearsing* can be considered as being low risk; employing students' *personal weaknesses* as peer assessment criteria are of greater risk,

particularly if students are inadequately prepared. A third focus relates to the potential of each attribute type in supporting and bringing about improvements to learning. The approaches towards establishing peer assessment criteria arising from the action research, which included those of encouraging students to be aware of their own personal weaknesses, either through self-selected or band-determined criteria, provided students with important learning opportunities.

This chapter has also addressed students' readiness and maturity in responding appropriately to the various types of personal attribute criteria that they were invited to formulate; this assumed an increasing importance in the final two years of the action research (especially in cycles 14 and 17). Finally, the increasing focus on motivation focus was discussed. This involved changing the types of attributes used as criteria for assessment during cycles 15-19. These cycles, with their focus firmly on individuals' personal weaknesses were intended to provide students with further motivation for improving and developing those areas.

The interventions which had the greatest consequence for each of these foci and, arguably, of having the maximum impact on the intra-peer assessments, involved the establishing of band-determined personal weaknesses in cycles 15, 16, 18 and 19. These interventions, more than any other, were an important catalyst towards the increasing focus illustrated in Table 4.1 (p.156). The following Table 4.2 summarises the impact that the use of band-determined personal weakness criteria had on the action research.

Table 4.2 Band determined personal weaknesses and their impact on the increasing focus of the action research

Areas of focus	**Increasing foci as a consequence of the use of band-determined personal weakness assessment criteria**
Increasing focus on individual	Criteria based on an individual's personal weaknesses
Increasing risk level	Significant risks to interpersonal relationships in asking individuals to agree to their band deciding his or her personal weaknesses and in using these as intra-peer assessment criteria
Increasing potential to support learning	Focus and band identification of each band member's personal weaknesses allows scope for improvements and development of these in rehearsing
Motivational focus	Increased motivation focus caused through individuals receiving feedback relating to their personal weaknesses and knowing that these will be used as intra-peer assessment leads to efforts to improve these.
Increasing focus on maturity or 'readiness'	Requires maturity and 'readiness' to agree to allow bands to decide individuals' personal weaknesses.

The discussion and analysis of the various focus issues, arising from the rehearsing and performing cycles, are explored further in Chapter Five. They are revisited once more in Chapter Six, where their significance is apparent in the models that are proposed as an outcome of the enquiry.

Chapter Five: Qualitative and Quantitative Data Analysis

5.1. Interview data analysis.

5.2. Peer assessment data analysis.

5.3. Relationships between the data analyses

5: Qualitative and Quantitative Data Analysis

This chapter presents an analysis and interpretation of the data that was produced from the rehearsing and performing cycles. Each cycle generated an extensive range of information that, year by year, produced a considerable qualitative and quantitative data resource. This data helped to inform the interventions that were made in the cycles at the time, as well as the shape of the enquiry.

It was explained in Section 3.6 how the qualitative data (interviews with the students) and quantitative data (the marks that were awarded in the peer assessments) were collected and analysed. The purpose of this chapter is to consider the analyses of each of these two data types. Chapter Five, therefore, presents an analysis and interpretation of the qualitative and quantitative data, together with a synthesis of its findings. It culminates in the development of a number of conceptual and theoretical explanations and statements founded upon the substantive categories that are indicated from the analyses and interpretation of the peer assessment data.[1] Key points arising from this analysis are examined further in Chapter Six and Chapter Seven. They include: building trust, promoting honesty, developing confidence, awareness, self-knowledge and feedback. This chapter is organised into three sections, which considers the qualitative and quantitative data analysis as follows

5.1. Interview data analysis.
5.2. Peer assessment data analysis.
5.3. Relationships between the data analyses.

These sections expand the analytical thread that commenced in the 'rehearsing and performance cycles' of the previous chapter (i.e. personal attributes and process features of intra-peer assessment). The next chapter develops the analysis and interpretation that is discussed here, by proposing a new process model of intrapeer assessment.

The first section, *Interview data analysis,* presents the analysis of the sixteen individual semi-structured interviews that were conducted during 2002-2005. Grounded theory procedures were adapted to analyse this data, (as explained in Section 3.6.3 of Chapter Three) the analysis of open, axial and selective coding processes that were applied is

[1] A number of examples that illustrate, in more detail, the analytical procedures that were applied to the data can be found in the *Appendix* (pp.281-288).

presented in this section. This section also contains the analysis of the process features line of enquiry (including private or collaborative marking and when in the process this should be conducted). Concluding the analysis is a diagrammatic representation that integrates its key findings into a central focus.

The second section, *Peer assessment data,* presents the analysis of the marks that were awarded for the intra-peer and inter-peer assessments during 2000-2005. This analysis culminates in a number of propositional claims.

Both data types are, however, interrelated across a number of areas. For example, students responded to the peer assessments marks that were awarded to them in their interviews. Free riders and shooting stars (as defined in Section 1.3.6) were also among the volunteers who were interviewed. Responses to the personal attribute criteria that students used appear in the interviews and also in the quantitative data analysis, in terms of the extent to which bands agreed over their marks for personal weaknesses, personal strengths and group attributes. These relationships will be examined in Section 5.3.

5.1 Interview data analysis

I conducted sixteen semi-structured interviews with individual Performance Management students and transcribed these also. An example of a transcript (p.282) and the analysis procedures (p.286) can be found in the Appendix. A summary of the interviews conducted in each cohort during 2002-2005 appears in Table 5.1.

Table 5.1. Interviews with Performance Management students 2002-2005

Year	Student interviewees						Total	Cohort	Sample
2002-2003	SA	OD	HA	MR	MK	WK	6	25	24%
2003-2004	NL	EJ	SH	DS	DC		5	19	26%
2004-2005	OM	YL	TP	EN	HS		5	12	42%

Student interviewees were denoted using letters that concealed their identity (SA, OD, etc.), in order to preserve their anonymity in this research. Despite interviewees being selected on a voluntary basis, the sample was representative of the cohorts as a whole. For example, the instrumental mix was typical and proportionate for a band line-up and comprised the following: six vocalists (two of whom also played guitar and keyboard), four guitarists (two of whom also sang), three keyboard players (two of whom also sang), two electric bass players and one drummer. In terms of their age profile, five

interviewees (MR, NL, EJ, OM, EN) were mature students; this reflects the proportion of mature students in the population (described in Section 3.1). Nine males and seven females were interviewed and, as such, female interviewees were slightly over-representative of the population as a whole. Three interviewees (WK, TP and EN), resulting from a number of intra-peer assessments were 'free-riders,' as defined by the criteria explained in Section 2.3.6. A balance, however, was provided by the participation of interviewees SA and NL who appeared as 'shooting stars' in certain intra-peer assessments, as defined in Section 2.3.6.

It was hoped that the interviews would be conducted at the very end of each academic year, following the final *Performance Management* band performances. However, these performances were always the final assessment of the examination period and, following these, many students left shortly afterwards for vacation work or to return home. Although I was able to conduct a few interviews at the very end of the summer term, it was more usual for these to take place during the following autumn term.

The interviews informed the interventions in the cycles and the direction of the enquiry.[2] The analysis of the interviews, however, was usually conducted at a later stage. Transcribing the interviews myself helped me acquire a greater familiarity with the data, which proved very beneficial during the coding analysis. Table 5.2 is an overview of the categories that were produced from the analysis of the interview data, which were coded into the following thematic categories.

Table 5.2. Overview of the categories produced from the interview data analysis

Thematic categories	**Coding procedure**
16 interviews produced 233 thematic categories	Open coding
8 core categories	Axial coding
8 core categories broken down into 22 subcategories	
22 subcategories	Selective coding
4 areas emergent from the 22 subcategories representing the central foci of the interview data analysis	

Analysing the data involved coding the responses given by the students in each interview into themes. This open coding produced 233 identifiable themes, or 'open coding' categories, from the sixteen interviews. These themes were further analysed in order to identify and classify core categories.

[2] Illustrative examples appear in Appendix 7 (p289).

A total of eight core categories were identified and are listed in Table 5.3. An illustration of the data analysis and coding procedures that I employed appears in Appendix sections five and six. These categories arose from grouping the open coding categories by core thematic relationships. Some of these categories (*Status of marks*, for example) might have been anticipated as a consequence of the focused questions that were included during the interviews.[3]

Table 5.3 Axial coding: eight core categories

Bias
Collaborative and private assessment
Impact upon peer assessors
Marking agreement
Personal attributes amenable to development
Responses to receiving feedback
Selection of personal attributes
Status of marks

In order to reveal possible relationships within each of these core categories, they were subjected to further detailed analysis, leading to a re-classification into sub-categories. This procedure for identifying core- and sub-categories arising from the interview data can be likened to *axial coding*. Sub-categories emerged from the open coding analysis by examining similarities between these and identifying whether they were also found among different students. Sub-categories were considered where similar categories were found in the open coding thematic analysis of at least three different interviewees. The basis on which sub-categories arising from the data were identified was explained in Section 3.6.3.

The analysis produced a total of twenty-two sub-categories identified by this process and their relationship with the core categories is indicated in Table 5.4. The names given to each sub-category were derived from words and phrases that students themselves articulated in their interviews. Naming the categories in this way would, I thought, be in keeping with the spirit of the analysis being *grounded* in the action research interview data. A measure of the strength of expression arising from the interviews for each category is also indicated in Table 5.4 (p.164). These measurement boundaries, described as 'strong', 'moderate' and 'weak' were, as explained in Section 3.6.3, derived from the proportion of interviewees whose responses are represented by that category, together with the number of open coded responses attributed to it.

[3] This is illustrated in Appendix 4 (p.278).

Students who expressed differing views were coded into other categories. Thus, where the sub-category *Marks should count* attracted a 'moderate' strength of response or comments, the opposing view, represented by the sub-category *Marks should not count*, attracted a 'weak' strength of comments or response. Naturally, students did not always articulate their views on every aspect of their peer assessment experiences in the interviews.

Table 5.4 Axial coding: core categories and sub-categories

Core categories	Sub-categories and strength of response
Bias	1.Friendship marking is a problem (weak*)
Collaborative and private assessment	2.Preference to mark privately (strong*) 3.Private marking is more honest (moderate*) 4.Prefer to mark collaboratively (weak)
Impact upon peer assessors	5.Honest feedback sought with everyone, including personal friends (strong)
Marking agreement	6.Marking either before or after knowing performance grade does not influence the intrapeer assessment (strong) 7.Marking is best before performance (moderate) 8.Marking is reasonably accurate and fair (strong)
Personal attributes amenable to development	9.Individual attributes are amenable to development (moderate) 10.Intra-peer assessment assists students' confidence (moderate) 11.Group attributes are amenable to development (moderate) 12.Problems inhibiting the development of personal attributes (moderate) 13.Thinking about peer assessment during rehearsals (moderate)
Responses to receiving feedback	14.Intra-peer assessment is good as it reveals what people think of you (moderate) 15.Intra-peer assessment gives you a check of what you are actually doing (moderate) 16.It teaches you to try harder, to do better (moderate) 17.It reveals your weaknesses to enable you to understand how to improve (moderate) 18.Need to emphasise constructive criticism (moderate)
Selection of personal attributes	19.Band members should choose individuals' attributes because of their knowledge of band members' rehearsing qualities (moderate) 20.Perceptions of self-weaknesses are not what others perceive as weaknesses (weak)
Status of marks	21.Marks should count (moderate) 22.Marks should not count (weak)
*Weak = similar open coding responses from three separate interviewees that additionally attracted several comments during their interviews. *Moderate = similar open coding responses from between four and seven separate interviewees that additionally attracted several comments during their interviews. *Strong = similar open coding responses from between eight or more separate interviewees that additionally attracted several comments during their interviews.	

5.1.1 Interpreting the axial core categories and sub-categories

How might the eight core categories and their twenty-two subcategories be interpreted, in terms of what the students were saying about their experiences of peer assessment? Although several of these questions, explained in Section 3.6.1, addressed learning experiences (questions: b, e, f, g), there were others that were not obviously so (questions a, c, d, h). However, the very large proportion of categories that may be thought to be related to students' learning experiences in peer assessment was still something of a surprise. For example, the sub-categories *Impact upon the peer assessors, Personal attributes amenable to development, Responses to receiving feedback, Selection of personal attributes,* are all suggestive of the learning activities or experiences. There is a key point to be made here: the analysis indicates that many interviewees' responses, which arose from their engagement with peer assessment, were related to learning. This resonates deeply with constructivist theories of learning and this aspect is revisited in Chapter Seven.

In returning to Table 5.4 (p.164), the next stage of analysis involved the identification and development of a central focus, representing the principle concerns of the students, which integrate the axial coding categories. This is a procedure that can be likened to *selective coding* in grounded theory methods. I shall briefly consider each of these sub-categories in turn before returning to offer interpretations of the analysis and what it means in the next section.

- *Friendship marking is a problem*

Although there were many open coding categories relating to the core-category of *Bias* (for example, comprising comments such as "I do feel that some people will give their friends more, a better grade" –OD), the only sub-category that could be formed was *Friendship marking is a problem.* This then, was the single strongest theme, deriving from *Bias,* which troubled students (for example, "Personal feelings get in the way sometimes, and people will purposely mark low out of spite" –EJ; "Unfair marking can lead to retaliation" –NL).

- *Prefer to mark collaboratively*

This sub-category is related to the process features line of enquiry described in the previous chapter. Although the analysis produced the sub-category *Prefer to mark*

collaboratively, (because, for example "you can feed off each other" –DC) which indicated a 'weak' strength of response, it also produced two other sub-categories in which others expressed an opposite view (for example, *Prefer to mark privately* and *Private marking is more honest*). There were 'strong' responses that expressed a preference overall, however, for intra-peer assessment to be conducted in private, with several students explaining this preference in terms of their belief of it contributing towards a more honest process.

- *Honest feedback sought with everyone, including personal friends*

This was a sub-category that emerged from the core-category *Impact upon the peer assessors.* That this sub-category specifically linked honest feedback with friends is interesting, as it suggests some difficulties were experienced in that area. As a whole, seeking honest feedback was an issue of some importance, being raised by eight of the sixteen interviewees. Those students clearly expressed a desire to be told the truth from their peers, which from their closest friends might not have been an easy experience.

- *Marking either before or after knowing performance grade does not influence the intra-peer assessment*

This sub-category, like the previous is also related to the process features line of enquiry. A total of nine interviewees felt that knowing the band's performance mark would not influence how they marked their peers in rehearsing. The implication is that the success or otherwise of bands' performances, as measured by their band mark, was irrelevant to individuals in how they awarded their intra-peer assessments. Other students, however, did express a moderate preference for *Marking before the performance* and being a sub-category such views (for example, "intra-peer marking best without knowing performance mark" –WK) have weight. This suggests that conducting intra-peer assessments before the band performs can help to prevent individuals being influenced inappropriately in their marking. There is ambiguity here: one sub-category reflects a view that is broadly neutral; the other specifically advocates when intra-peer assessment ought to be conducted.

- *Marking is reasonably accurate and fair*

Although there was a positive response to the accuracy and fairness of the marks that were awarded, it says little about how or why the marking was so adjudged. A separate sub-category that might have explained this, unfortunately, did not emerge from the

data. Does the absence of this indicate students' inability to articulate the difficulties involved in making sophisticated judgements? An examination of several open coding categories does, however, suggest that some students felt that the marks were fair due to their band providing honest assessments about themselves and about the other band members. A concern for honesty among students, however, was also apparent and overlapping in sub-categories elsewhere (for example, *Honest feedback* and *Private marking is more honest).* Honesty, a key theme to emerge in the coding, was evident in many instances of students' peer assessment experiences, and this subject is revisited elsewhere in this as well as in the following chapters.

- *Thinking about peer assessment during rehearsals*

I certainly hoped that students would be thinking about peer assessment at some point during their rehearsing and the eventual activity of awarding marks ("subconsciously always thinking about the peer assessing" –TP; "...thinking about qualities while rehearsing: being flexible, along with other attributes" –DC). Like several sub-categories arising from the analysis, *Thinking about peer assessment during rehearsals* was also one that seemed to overlap across others (for example, sub-category 14, *Intra-peer assessment is good as it reveals what people think of you).* Thinking about peer assessment in the context of rehearsing, being expressed by six interviewees attracted a 'moderate' strength of response. It is a sub-category that is perhaps consequential on using personal attributes as criteria for intra-peer assessment. It suggests that knowing that their rehearsal activities would be assessed using their personal attribute criteria caused students to think and reflect on this experience, and this may have influenced their behaviour. It might suggest, therefore, that students were engaged in thinking that went beyond cognition, and, as a sub-category, is of particular importance to learning.

- *Intra-peer assessment gives you a check of what you are actually doing*

Similar to some others, this sub-category related to feedback and represents the coded responses of four of the interviewees. It was similar, for example, with *Intra-peer assessment is good as it reveals what people think of you.* Together, these two related sub-categories represented many of the profound experiences that were often expressed by interviewees ("it lets people know what they think of you as well as what you think of yourself" –OD) as a consequence of their involvement in intra-peer assessment.

- *Perceptions of self-weaknesses are not what others perceive as weaknesses,*

This sub-category indicates that some interviewees (for example, HA, SA and NL) were surprised by the *personal weakness* attributes that their band had formulated for them. There was a realisation emerging from their surprise that the band had perceived something about them of which they were previously unaware, but had now acknowledged. The experience described by some when receiving these, was that of surprise –even revelation ("my [self-selected] weaknesses – they're not what the other people would perceive as my weaknesses" -SA)

- Intra-peer assessment involving personal attributes criteria provides students with feedback, but is also, articulated by NL "an exceptional way of getting feedback". Three sub-categories emerged which, together, suggested that students responded positively to using individual attribute criteria, which they linked with feedback. Two sub-categories suggested that interviewees' responses to the feedback from intra-peer assessment were comprehended in different ways. Thus, *It reveals your weaknesses and once you know these you are fine* suggests a response that was characterised by a revelatory experience, followed by realisation, acceptance and tacit knowledge of what is required. *It teaches you to do better and try harder* suggests, how ever, a motivational response. The third sub-category expressed interviewees' firm views (for example "former band members who rehearsed for a performance with you, so they could choose these attributes for you, that you can work on.... so they know you..." –SA) about who was most appropriate to determine individual attribute criteria: *Band members should choose individuals' attributes because of their knowledge of band members' rehearsing qualities.*

- Sub-categories emerged from interviewees' responses to using personal attribute criteria that they considered supported their learning. These are represented by the sub-categories: *Individual attributes are amenable to development* and *Group attributes* [4] *are amenable to development.* Each of these sub-categories arose through interviewees identifying and explaining particular attributes that they considered to have developed as a consequence of intra-peer assessment ("punctuality and reliability improved" –MK; "I learned more about working with different teams of people" –OM; "using peer assessment encourages commitment" -HS).

[4] Group-agreed attributes previously explained.

Even after excluding 'confidence' which became a sub-category in itself, a total of seven students cited various individual attributes that they considered had developed. *Intra-peer assessment assists students' confidence* was a sub-category that appeared to have particular significance. Of all the personal attributes that were used as intra-peer assessment criteria, it was *Confidence* that attracted the greatest strength of expression as the attribute that had developed the most as a consequence of peer assessment.

- A sub-category *Problems inhibiting the development of personal attributes* emerged from the analysis. This sub-category, arising from the responses of seven interviewees, was formed from a range of difficulties that were raised by interviewees ("contribution to ideas or leadership difficult if there is a dominant member" –HS; "with 'creative input' I have to say that in this particular band I was a backing singer, so my 'creative input' for the whole process wouldn't have been too high; so it is a defence really, of how my marks have gone down" –SA; "absence from rehearsals is a problem" –MK). The analysis did not identify any one single problem directly related to developing interviewees' personal attributes more than any other. Taken together, therefore, this sub-category suggests a variety of difficulties that related to individuals' personal circumstances, than to a single common difficulty.

- The *Need to emphasise constructive criticism* emerged as another sub-category that was related to feedback. Four interviewees emphasised their desire for constructive criticism ("I need constructive criticism, not a stab in the back" –WK). This suggests that their experiences of receiving feedback were possibly less than satisfactory. As explained earlier, face-to-face feedback might have been given to the individual by their band during the activity of presenting the band-determined personal weaknesses assessment criteria. Feedback was also given when students received their intra-peer marks. This sub-category suggests that the emphasis for each of these feedback methods should be of supporting learners in providing constructive criticism.

- Elsewhere, the analysis identified key feedback issues that involved the giving and receiving of marks and students were divided on whether marks should contribute towards grades (core category *Status of marks*). Of those that expressed a firm opinion, the majority thought that marks should indeed contribute to summative grading, some citing reasons, for example, that "students take it more seriously" -TP.

At this point it may be helpful to provide a brief summary before continuing with the analysis. A total of sixteen individual semi-structured interviews were conducted with Performance Management students during 2002-2005. Grounded theory was adapted to analyse the student interview data. An *open coding* process yielded a total of eight core conceptual categories. These were then analysed further in order to identify possible sub-categories within the core categories, a procedure similar to *axial coding*. A total of twenty-two sub-categories emerged from the data analysis. A number of key concerns arising from the data and which appeared to overlap and encompass several sub-categories began to emerge, for example: one of these concerned feedback; honesty was another. The next part of the analysis, involving selective coding procedures, attempts to integrate the key concerns into a central focus of the data. This selective coding also uses a number of illustrations in order to more easily explain, in visual terms, the central focus of the data and the emergent meanings and conclusions.

Four areas emerge from subcategories that suggest a focus for the data, which are: self-knowledge; feedback; honesty; confidence. These are the substantive categories derived from the grounded theory analysis and, as illustrated in Figure 5.1, encompass all the sub-categories of the data. They provide a central integrative focus that represents the key concerns of the interviewees' experiences of peer assessment. It is interesting to note that none of these areas directly equates with the focussed questions, as described in Section 3.6.1. Their integration characteristics will be discussed after the analysis and interpretation of the data. Each of these four foci will be considered in turn, beginning with self-knowledge.

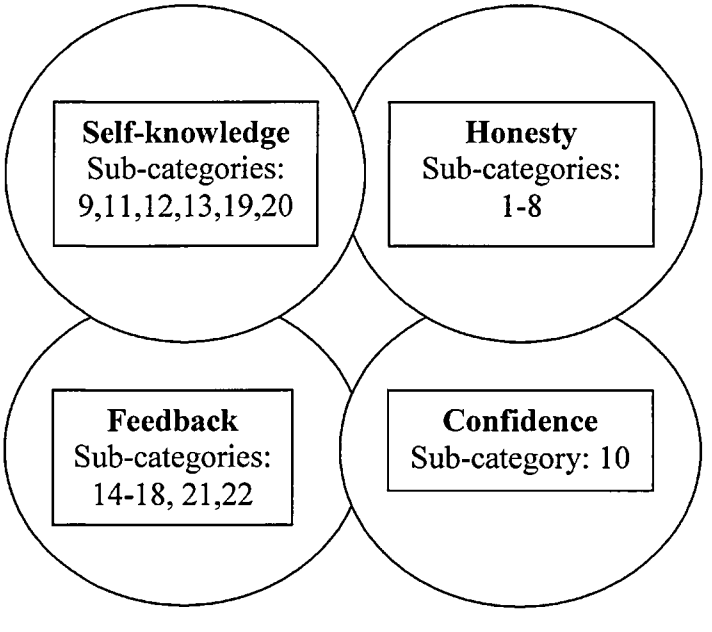


Figure 5.1 Selective coding foci of interview

The focus on self-knowledge

There were many sub-categories that suggested a focus on self-knowledge or self awareness arising from interviewees' engagement with personal attribute criteria. Six categories revealed a relationship with this area. Through the process that involved students identifying and thinking about themselves in terms of the personal attribute criteria, three observations can be made here. Firstly, the interviewees possessed knowledge of themselves through engagement with their personal attributes criteria which may have led to thought, action and change in terms of developing their learning during rehearsing (indicated by 9,11,12 and 13). Secondly, being able to access that knowledge about themselves (indicated by 19 and 20) could be fundamental to learning. It was conducted through two mechanisms: self-selected attributes usually self-selected personal weaknesses or band determined personal weaknesses.Therefore, self-knowledge arose from these activities: developing self awareness and, knowing about themselves through their peers.

A third observation involves the nature of each source of knowledge. Knowledge arising from interviewees' self-awareness ("it made me look a lot more at myself" – YL), as a consequence of thinking about their personal attribute criteria ("subconsciously, I was always thinking about them two things I needed to improve on" –TP), becomes important if intra-peer assessment assists in access and externalisation (i.e.enabling such knowledge to be available to peers). Knowledge of oneself that is derived from one's peers is of a different status. It is different because it is created from the collective experiences of one's band members and it exists externally to the student. The context in which such knowledge is created especially resonates with social constructivist theory.

Accessing such knowledge was regarded as being of critical importance by interviewees (indicated by 19) and the strength of expression that they attached to the significance of accessing this externally created band knowledge about themselves is emphasised by 19 and 20.The theoretical underpinning, relating to the acquisition of such knowledge through intra-peer assessment, will be discussed in the following chapter. The focus on self knowledge is summarised in Figure 5.2 (p.172).

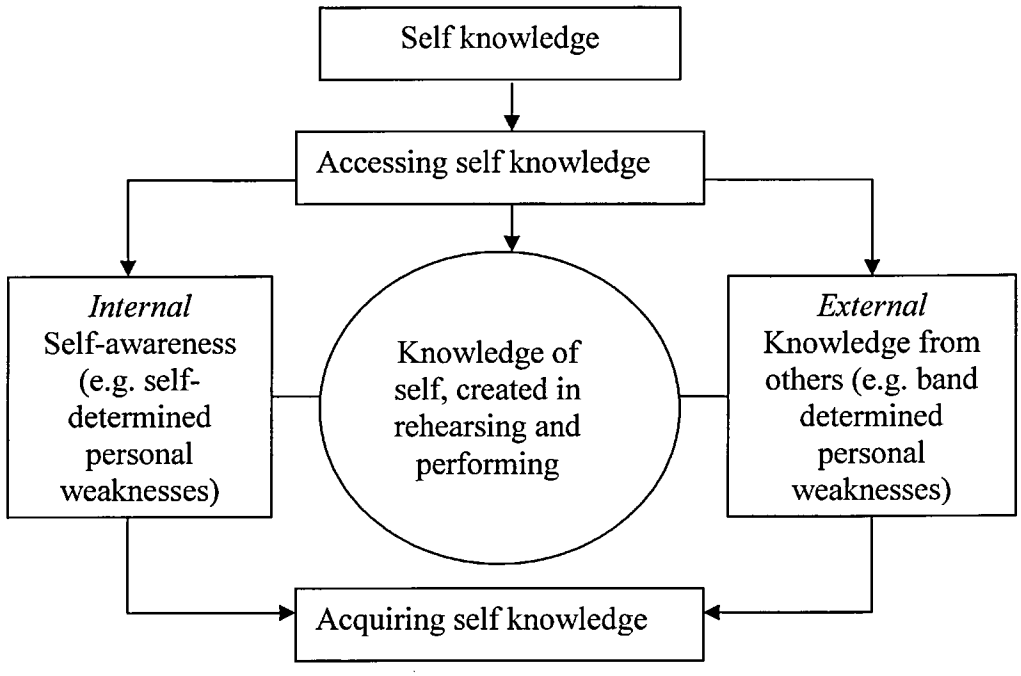


Figure 5.2 The focus on self knowledge

The focus on feedback

This is the second selective coding focus identified in Figure.5.1 (p.170). Many students considered the intra-peer assessments as, in the words of one interviewee, "*an exceptional way of getting feedback*", allowing individuals to learn from ongoing peer feedback from their band members about their personal attributes. There were many subcategories that suggested the importance of feedback. These were indicated in the axial coding categories *It is good as it reveals what people think of you* and *It gives you a check of what you are actually doing*. Many interviewee comments about feedback arose from two sources: receiving band-determined personal weakness attributes and, receiving the intra-peer marks that their band awarded to them. There were many comments involving the former: feedback arising from receiving their band-determined personal weaknesses. The data arising from interviewees' responses to this feedback indicates that it falls into two kinds: revelatory feedback and motivational feedback.

(1) Revelatory

For some, in receiving the kind of feedback experienced where their band reveals their personal weaknesses, this can be revelatory and chastening. Interviewees' responses (involving 14, 15, and 17) indicate that students quickly move through a number of stages when comprehending this feedback. One of these can be interpreted as 'realisation' ("you have to listen and respect someone willing and brave enough to tell

you your faults" –NL). The next stage may involve 'acceptance' ("I agree with the attributes chosen for me by my band" –YL) and onwards towards the acquisition of a 'tacit' kind of knowledge' ("it enables you to understand how to improve"- EJ). Such revelatory experiences, and how these might support students in the acquisition of self-knowledge, are examined in the next chapter.

(2) Motivational

For others, their intra-peer assessment feedback comments suggested a motivational response. For example, *constructive* feedback from peers was desired (16, 18) and receiving it was expressed in terms of *learning* followed by *action* ("it teaches you to try harder, do better"-OD, HA, NL, HS; "I made a conscious effort to work on my attributes" –EN). This comment is interesting in that it shows an acceptance of the intrapeer assessment framework. Most interviewees, who thought that marks should count towards grades, expressed their view in terms of motivation ("without the marking we would not put the same effort in improving these areas" –DS, EN; "it's taken more seriously if marks count" –TP; "marks should count because practicing is a big part of it" –HJ, DC).

Being able to respond appropriately to such feedback, whether in receiving *personal weaknesses* attributes that the band has determined, or in receiving their intrapeer assessment of these, attracted some comments. One student, for example, believed that as a pre-requisite for allowing bands to decide individuals' personal weaknesses "you have got to be able to take constructive criticism for others to decide your attributes" - EJ; another discovered that it was "hard to work with people who cannot accept criticism" –OM. Similarly, interviewees commented upon the "need to emphasis constructive criticism… [rather than a] …stab in the back" – WK. Similarly, when receiving disappointing marks "if criticism makes you take stock, then peer assessment is OK" –NL.

These responses (i.e.revelatory and motivational) arising from students' interviews, suggests that the knowledge that may arise from this type of feedback can support learning. Figure 5.3 (p.174) summarises the focus on feedback.

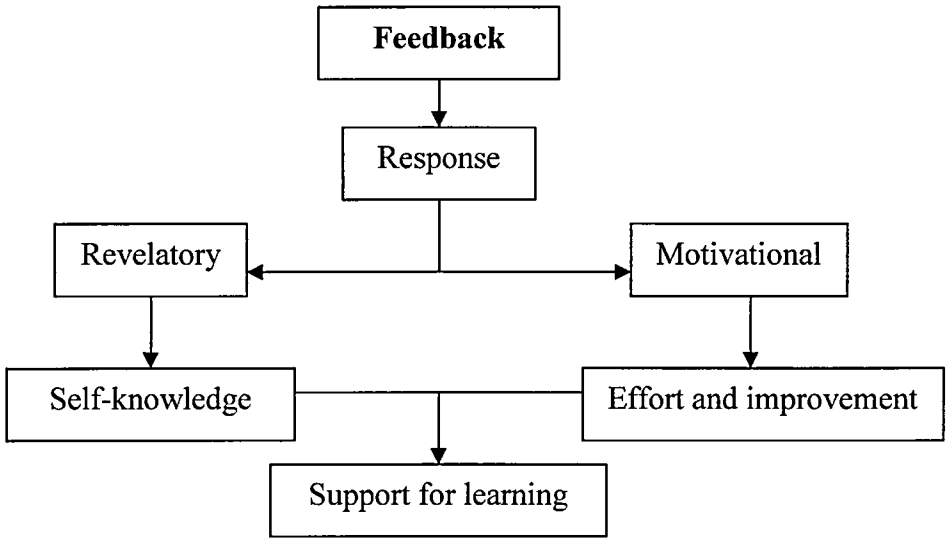


Figure 5.3 The focus on feedback

The focus on confidence

Although only one subcategory emerged that specifically highlighted confidence as a theme deriving from the interview data (10), this might have been consequential on the interpretation of the thematic coding. For example, it is not clear whether 'confidence' should be considered as a discrete category as it appears to be implicit in other categories also (9,11,20). Within category 10 alone, it was raised by no less than seven interviewees. The coding included comments about confidence being displayed by other band members as well as responses about individuals' inner confidence.

Indeed, the data suggest that the development of students' confidence was manifest on at least two levels. Some statements indicated 'confidence' in a context of its particular (and frequent) usage as a personal attribute criteria, typically as a *personal weakness* attribute (i.e. *lack of confidence*) that interviewees stated had improved as a consequence of intra-peer assessment ("I suppose in the rehearsal process in the past, I've always been the one to kind of take a back seat and kind of just sit and do whatever I'm told to do, if you know what I mean. So, because they put me in the position of, you know, deciding this, that and other, I've learnt to be more assertive" –DS). Indeed, it was this *confidence* attribute, more than any other, that was cited by the interviewees in terms of experiencing changes in their behaviour ("[peer assessment]... allows you to notice confident people ... [and]...it allows you to notice shy people" –OM). The prominence of this attribute over others also extended to its usage as a criterion for individual attributes throughout the rehearsing and performing cycles. It was the most

frequently formulated individual attribute over all others (comprising in excess of one hundred other different individual attribute descriptors used in the enquiry).

Another level comprised general expressions of increased confidence arising from interviewees' engagement in the intra-peer assessment process rather than specific elements of it ("It makes people stronger" –MR; "It made me look a lot more at myself, and looking at their confidence like, with mine" –YL; "You definitely notice people who are confident on their instruments, put forward, put their ideas forward more than people who tend to maybe not have had as much experience of playing with people; they like to sit back a little bit" – OM).

In conclusion, the totality of these arguments, the analysis, and its interpretation is cumulative and compelling; they are supportive towards the integration of 'confidence' as a substantive category. The focus on confidence is illustrated by Figure 5.4.

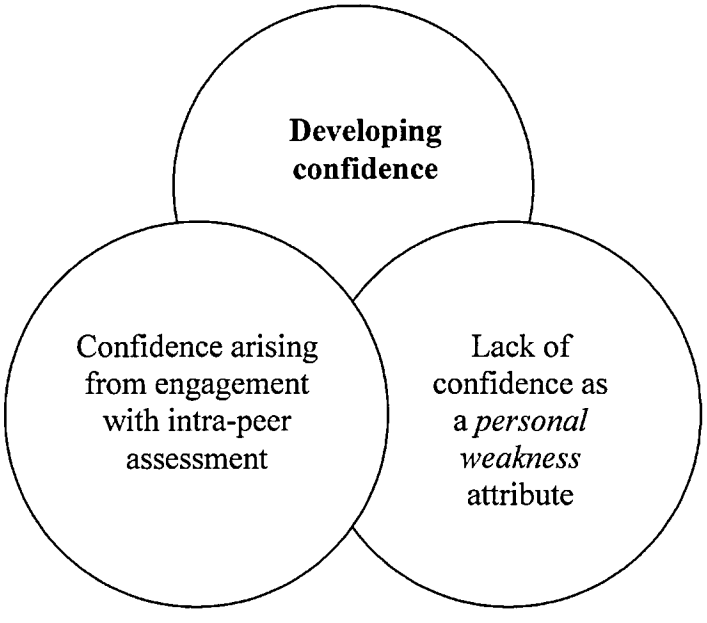


Figure 5.4 The focus on confidence

The focus on honesty

There were many subcategories that indicated a focus on honesty (1, 2, 3, 5, 6, 7, 8) including four falling into the 'strong' boundary and two belonging to the 'moderate' boundary. In view of this the focus on honesty was of significance in the selective coding process. Interviewees' comments that referred to honesty suggested two concerns: one arising from issues of trust in peer assessment processes and the other

arising from band members' interactions and encounters with each other. In Figure.5.5, these are interpreted and described as being 'the trust dimension' and 'the social dimension' respectively.

Interviewees' responses relating to the trust dimension were mainly as a consequence of the focused questioning within the 'process features' area. Many expressed the view that intra-peer assessment should be conducted in private, as they felt it would lead to marking that was more honest (2, 3). Nevertheless, conducting intra-peer assessment in private, can lead to increased opportunities for some to mark unfairly, in view of the anonymity of the process. However, even in an entirely open or collaborative marking setting, it is "*difficult to try to influence people to be impartial if they are determined to give low marks*"-EJ.

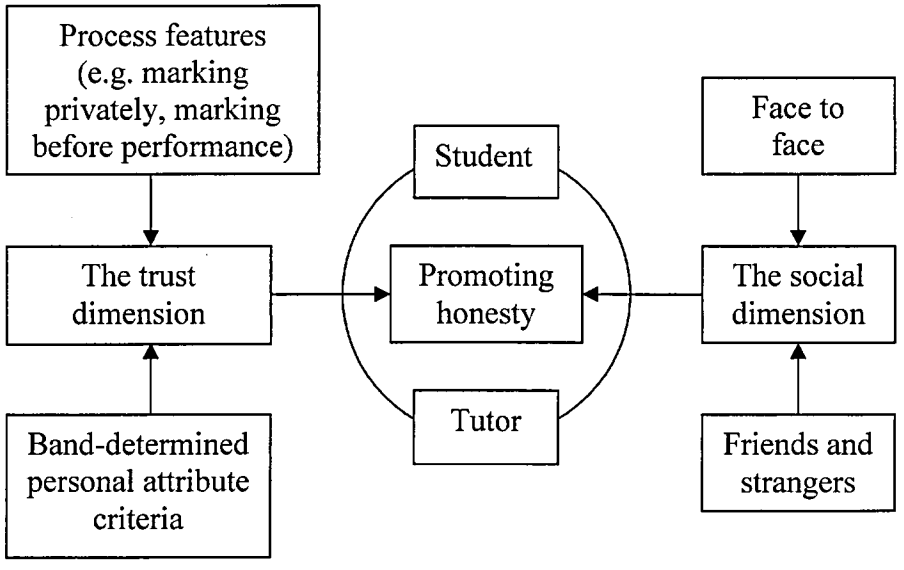


Figure 5.5 The focus on honesty

One third of interviewees favoured intra-peer assessment to be conducted prior to the performance. This suggests that not all students felt they could trust each other to be inappropriately influenced by the success or otherwise of their on-stage performance (7). A number of students favoured band-determined individual attribute criteria specifically on the grounds that it was more honest than self-determined *personal weaknesses* ("band members should choose, because they are more honest at identifying your weaknesses" –SA). Comments that related to 'the social dimension' displayed varying layers of concerns and anxieties. Of these, two areas were particularly apparent:

firstly, students' apprehension in face-to-face encounters, in their reporting, for example, of "difficulties in telling the truth face-to-face, so written feedback is good" – SA, DS; "written feedback avoids confrontations" – HA, SA); secondly, how students felt about assessing friends and strangers ("it's tough assessing your friends" –OD;" "I try to be honest with my friends" -HA, MR"; "some give friends better marks" –OD; "am more critical with strangers" –OD)

5.2 Peer assessment data

The analysis and interpretation of the previous interview data, identified four key areas: acquiring self-knowledge; feedback and supporting learning; developing confidence; promoting honesty. This section presents the analysis of the peer assessment marking data during the period 2000-2005. As described previously in Section 3.6.5, the quantitative data (with a population size of 170 and a 100% sample) comprised an extensive and detailed corpus, forming a useful informational background (Guba & Lincoln, 1994) to the interview data and indeed to this naturalistic enquiry as a whole.

This data also share a relationship with the previous interview data in the analysis of students' responses, particularly in the core categories *Marking agreement* and *Status of marks*. Examples of this procedure and analysis appear in Section 8 of the Appendix.

Table 5.5 provides an overview of the procedures that I adapted from grounded theory in order to analyse the peer assessment data. These are discussed below and examples of the analytical procedures that were applied to the quantitative data also appear in the Appendix 8 (p.291).

Table 5.5 Overview of the analytical procedures

Type of coding	Techniques used	Primary data	Secondary data	Tertiary data
Open	Quantitative coding (i.e. high, fair, low)	Intra-peer marking data for each year	Statistically treated assessment data, e.g. standard deviation; normalised difference coefficients;	Uses primary and secondary data
	Comparative coding (i.e. higher than, lower than, equal to)			
Axial	Combining the respective sub-categories to produce a final total		Combining the respective sub-categories to produce a final total	Correlation of primary and secondary axial data (analysis summary matrices)
Selective	Diagrammatic representation of the central focus of the data analysis			

At this point, reference may be made to pp.127-130, which offer a rationale for adopting grounded theory principles in order to analyse the quantitative peer assessment data. Of relevance to this section are a number of key points that were made there and elsewhere in Chapter Three; these are summarised as follows:

- Grounded theory techniques are appropriate to naturalistic enquiries (Cohen *et al*, 2000; Piantinida & Garman, 1999);
- Although associated with studies involving qualitative data, grounded theory is, nevertheless, considered equally suitable for quantitative contexts (Glaser & Strauss 1967; Strauss & Corbin, 1998; Robson, 2002);
- Adapting grounded theory techniques for analysing the peer assessment data in this study are founded on its principles of 'simplicity' and 'fitness for purpose' (Cohen *et al*, 2000);
- Such principles informed the decision to select simple, low-powered statistical tests in order to provide analytical information relevant to the purpose of enquiry and to strive to maintain the links between the assessment data and the context in which it occurs (Miles & Huberman, 1984);
- The coding techniques employed should arise from the data, be fit for the purposes of the investigation and have meaning for its lines of enquiry;
- The assessment data suggested two simple coding procedures that would satisfy the above criteria: one which would code the numerical measurement data into basic categories described as 'quantitative coding'; the other which would compare marks and the categories, described as 'comparative coding'.

Quantitative coding

Three basic categories arising from the numerical measurement data set that would be relevant to the enquiry were suggested by using the coding categories of 'high', 'fair' and 'low'. These categories appeared to satisfy the 'fitness for purpose' criterion in that they might reveal interesting information about shooting stars and free riders in particular. Moreover, as the coding fell into three categories, this allowed propositions to be made that were supported by a large data corpus arising from one of the category trio.

Analysis involving three categories only, with which to code the numerical measurement data, supported grounded theory analysis principles concerning

'simplicity' of procedure (Cohen *et al*, 2000). This provided sufficient discrimination in order to give meaning to the data while avoiding the spurious validity that may have arisen from a more fragmented suite of categories.

Comparative coding

A similar procedure was adopted in order to compare marks and categories that arose from the analysis. The data quite simply suggested the three-way coding categories of 'greater than', 'equal to' and 'less than' as explained on page 126. These categories appeared to satisfy the 'fitness for purpose' criterion in that they might reveal interesting information when, for example, comparing students' consistency of agreement between band members' group attributes and personal weaknesses. A similar argument to that appearing under the previous 'quantitative coding', supports the simplicity, relevance and validity of this procedure, although similar cautions, raised about whether or not it provided sufficient discrimination of the data, were applicable.

Examples of these analytical procedures:

For data that required 'quantitative coding' (for example, measuring consistency of agreement), I simply coded the marking data according to three categories: high, fair or low. 'High', 'fair' or 'low' was defined by measuring the variance of the raw marks.

Example.

The consistency of agreement in assessing 3 personal weaknesses (measured by a variance of 1 or less) would be coded 'high'.
The consistency of agreement in assessing 3 personal weaknesses (measured by a variance of +1 or – 1) would be coded 'fair'.
The consistency of agreement in assessing 3 personal weaknesses (measured by a variance of >+ or – 1) would be coded 'low'.

For data that required a comparative coding (for example, comparing group and individual attributes marks) I similarly coded the data according to three categories: higher than, lower than, or equal to.

Example.

Group attributes assessed higher than individual attributes.
Group attributes assessed lower than individual attributes.

Group and individual attributes assessed equally.

These procedures are the equivalence of the 'open coding' stage in grounded theory. The open coding categories arising from each of the assessments during the year (i.e. *Christmas Party, Venues & Audiences, Decades*) were then combined in the respective sub-categories to produce a final total. This procedure can be likened to the axial coding stage in grounded theory. These totals were then represented diagrammatically displaying their inter-relationships, in order to provide a central focus of the combined coding for that academic year. This procedure can be likened to selective coding in grounded theory. Finally, a number of propositional claims, arising from the diagrammatic selective coding were identified.

A similar coding procedure was employed for the statistically treated *Secondary data*. *Secondary data* that suggested 'quantitative' coding was simply analysed according to three different measurement ranges that were appropriate to the data.

Example.
Personal weaknesses exhibit a normalised difference coefficient of 0.90 or higher.
Personal weaknesses exhibit a normalised difference coefficient of 0.75 or higher.
Personal weaknesses exhibit a normalised difference coefficient of less than 0.75.

A high-normalised difference coefficient (for example, 0.90 or higher) indicates a strong correlation between an individual's self-assessment and that of his/her band. Conversely, a low normalised difference coefficient (for example, less than 0.75) indicates a weak correlation between an individual's self-assessment and that of his/her band.

Secondary Data that suggested 'comparative' coding was, as before, simply analysed and placed in one of three possible categories accordingly.

Example.
Personal weaknesses have a smaller mean standard deviation than group attributes
Personal weaknesses and group attributes have same mean standard deviation
Personal weaknesses have a larger mean standard deviation than group attributes

The procedure thereafter was the same as for the analysis of the *Primary data*.

The coding and analysis procedure that was adopted for *Tertiary data* differs from that used for *Primary* or *Secondary data* coding. As *Tertiary data* comprised the already coded *Secondary* and *Primary data*, open coding was not necessary. Correlation analysis was conducted on the *Secondary* and *Primary data*, and this involved assembling an 'analysis summary matrix' comprising the sub-categories deriving from *Secondary* and *Primary* axial coding. Coding the analysis summary matrices is the equivalent of axial coding. The procedure of selective coding followed the same as before.

I adopted a pragmatic approach in deciding the criterion for what constituted a *Propositional claim*. For the *Primary* and *Secondary data*, I decided that if any one of the three coded categories (e.g., *High, Fair, Low* or, *Higher than, Equal to, Lower than)* accounted for at least 50% of the total for that coding category, then this was sufficient evidence for a propositional claim. I considered this as being a significant boundary; a total comprising one half would exceed that of even a combined coding total for the two remaining categories. For *Tertiary data*, propositional claims were all based upon the *selective coding* grounded upon the *analysis of quantitative comparisons* tables.

All propositions were supported by a significant amount of quantitative data. For example, an examination just of the very first 'axial' coding category for *Primary data* of 2003 reveals that *Consistency of intra-peer student marking agreement in the assessment of individual attributes (personal weaknesses)* is based upon a total of some 208 intra-peer assessments awarded by all of the bands involved.

The separate claims arising from the analysis of each of the five years of peer assessments were then examined. Many of the claims from any one particular year were repeated in other years. There were remarkably few claims that appeared to be contradictory. The analysis was completed through the amalgamation of each year's claims (illustrated in Table 5.6, p.183).

5.2.1 Summary of the quantitative assessment data

To summarise, a large amount of quantitative data was analysed using an adapted grounded theory approach. A number of propositional claims, arising from the quantitative assessment data, were made and the relative strength of each claim was defined in terms of being *Primary, Secondary* or *Tertiary.* Many claims were associated with students' group attributes and their individual attributes (*personal weaknesses* or *personal strengths*). Also emerging from the data analysis were a number of claims concerning students who received the highest intra-peer band assessments (described as *shooting stars*) and those who received the lowest (described as *free riders*). Many of the claims that were proposed, derived from quantitative data associated with consistency of marking agreement, dispersal of marks, and, comparisons of self-assessments with band members' assessments. Table 5.6 presents the claims arising from the analysis of the quantitative data.

Table 5.6 Peer assessment data 2000-2005: Propositional claims (primary, secondary and tertiary claims)

	Primary claims
1	Moderate correlation between students' self-assessments of their *personal weaknesses* and their band members assessment of these.
2	Moderate correlation between students' self-assessments of their personal attributes and their band members assessment of these.
3	A small dispersal of marks (standard deviation) occurs during the assessment of *personal weaknesses.*
4	A very small dispersal of marks (standard deviation) occurs during the assessment of *group attributes.*
5	The assessment of *group attributes* exhibits a smaller dispersal of marks (standard deviation) than for *personal weaknesses.*

	Secondary claims
1	Students exhibit a high consistency of agreement with each other when intra-peer assessing their band members' *personal weaknesses.*
2	Students exhibit a high consistency of agreement with each other when intra-peer assessing their band members' *group attributes.*
3	Students award higher marks for *group attributes* than for *personal weaknesses.*
4	Students exhibit a high consistency of agreement with each other when intra-peer assessing *shooting stars.*
5	Students exhibit a high consistency of agreement with each other when intra-peer assessing *free riders.*
6	Students award higher marks for *group attributes* of *shooting stars* than for *personal weaknesses.*
7	Students award equal marks for both *group attributes* and *personal weaknesses* of *free riders.*

	Tertiary Claims
1	A high consistency of intra-peer marking agreement correlates with a high band performance mark.
2	A low consistency of intra-peer marking agreement correlates with a low band performance mark.
3	Students exhibiting large standard deviations of their *group attributes* assessments also exhibit large standard deviations of their *individual attributes* assessments.
4	Students exhibiting small standard deviations of their *group attributes* assessments also exhibit small standard deviations of their *individual attributes* assessments.
5	Students exhibiting small standard deviations of their personal attributes overall, also exhibit small self-assessment variances.
6	Students who under-estimate their personal attributes in awarding themselves relatively low marks (compared with how their band members assess these attributes) usually receive good final mark.
7	Students who over-estimate their personal attributes in awarding themselves relatively high marks (compared with how their band members rate these attributes) usually receive poor final mark.
8	There is a correlation between individuals who over-estimate themselves in terms of self-assessment, and receiving a small individual normalised differential coefficient.
9	Individuals who receive a low final mark also tend to exhibit a large standard deviation for *group attributes*.
10	Individuals who receive a low final mark also tend to exhibit a large standard deviation for both *group* and *individual attributes*.
11	Individuals who receive a low final mark also tend to exhibit a small individual normalised difference coefficient.
12	Individuals who receive a high final mark also tend to exhibit a small standard deviation for both *group* and *individual attributes*.
13	Individuals who receive a high final mark also tend to exhibit a small standard deviation for their *group attributes*.
14	Individuals who receive a high final mark also tend to exhibit a small standard deviation for their *individual attributes*.
15	Bands receiving a high performance mark also tend to exhibit a small standard deviation mean for their *individual attributes*.
16	Band members, who collaborate in determining their individual marks, decide on an equal allocation of marks disregarding individual contributions (except for *free-riders*).
17	*Free riders* over estimate themselves when self-assessing their personal attributes.
18	Assessing *free riders* is problematic for bands in terms of their consistency of agreement.
19	*Shooting stars* under-estimate their personal attributes when self-assessing compared with their band members' assessments of them.
20	Band members exhibit a high level of assessment agreement when assessing the *group attributes* of *shooting stars*.

5.2.2 Analysis and interpretation

At this point, the discussion turns to the analysis and interpretation of the marks that students awarded each other in their assessments. This extensive information source was analysed using grounded theory that was especially adapted to the quantitative assessment data.

The previous section presented a number of propositional claims, arising from the grounded theory analysis, and the strength of each claim was defined using a hierarchical structure (i.e. *Primary*, *Secondary* or *Tertiary*) as previously explained in Section 3.6.6. Many of these claims were associated with students' group attributes and individual attributes (personal weaknesses or personal strengths). There were also claims arising from the analysis involving shooting stars and free riders.

I shall first consider the strongest propositional claims in the hierarchy: *Primary* claims. Students consistently agreed with each other within their bands when intrapeer assessing personal weaknesses, shooting stars and, to a lesser extent, free riders. Indeed the strength of agreement generally among students' assessments across all intrapeer personal attribute criteria was quite extraordinary. This, I believe, is extraordinary given the assumptions that were made when originally formulating the marking system. Individuals awarded marks, almost always, on the basis of a 1-5 marking scale, involving simple single descriptors (i.e.1=poor; 3=average; 5=excellent). As explained in Section 3.3.1.2, I assumed that there would be no shared conceptual understanding or consistent marking standard apparent among students when, for example, awarding a mark of 3. Nevertheless, I expected that each individual would apply his or her assessments consistently from student to student: a mark of 4 awarded to student WX would equate to a mark of 4 awarded to student YZ, for example. The strength of marking agreement that was found overall suggests that this expectation was well founded. The consistently strong agreement may also be explained through a common understanding among the students of the semantics of 'poor', 'average' and 'excellent' (and the meanings that they attached to marks of 1, 2, 3, 4 or 5). A further explanation for the strength of marking agreement, apparent among most bands, might also involve the development of a 'tacit' understanding of marking criteria among the students (the significance of 'tacit knowledge' was discussed in Chapter One).

The consistency of marking agreement within bands when assessing free riders dipped slightly below those exhibited for shooting stars; this raises questions. Why, for example, were bands somewhat less certain in their assessments of the personal attribute criteria of the weakest members of their band? Was it as a consequence of the particular personal attribute criteria on which free riders were assessed? It would seem unlikely that particular personal attribute criteria could have such causal effects. After all, free riders (and shooting stars) were only described as such *after* intra-peer assessment, whereas personal attribute criteria were decided beforehand. A more plausible explanation, however, may be found in the process through which bands perceive and discriminate levels of achievement in their members. For example, potential shooting stars may be regarded by their peers as excelling in rehearsals to such an extent that even their personal weaknesses appear more impressive than the personal strengths of others. If so, it is less surprising that the personal weakness attributes of shooting stars' attract such relatively high marks. An examination of the analysis arising from the intra-peer assessment marks for shooting stars supports such an explanation.

In comparing standard deviations, those of the free riders were slightly larger than for shooting stars. If so, then perhaps assessing personal attribute criteria of would-be free riders was a more complex process. Whereas most of the marks awarded for shooting stars coalesced towards the maximum mark of 5, resulting in small standard deviations, the marks that were awarded to free riders usually ranged from 1-3, which produced slightly larger standard deviations. Also, at the marking boundaries of 1 and 5, there were considerably more marks of 5 (the maximum) awarded to shooting stars (indicating a stronger measure of agreement) than there were marks of 1 (the minimum) awarded to free riders. Analyses of most intra-peer assessments reveal peers' reluctance in awarding the lowest possible mark (of 1) to would-be free riders. In summary, it may be concluded that students found it easier to recognise a student deserving of being awarded the highest marks for their personal attribute criteria, than deciding who deserved to be awarded the lowest. Or, putting it another way, how poor does one need to be during rehearsals in order to deserve being awarded the lowest possible mark of 1?

Group attributes attracted the highest assessment marks overall, exceeding those of *individual attributes based upon importance to rehearsing* and *personal weaknesses*. It was to be expected that *personal weakness* attributes (whether chosen by the individual

concerned or determined by their band) would be marked lower; after all, the *personal weaknesses* category should, by its description, attract the lowest marks of any.

A slightly different picture emerges regarding the assessments of would-be free riders. Although free riders were awarded poor marks in equal measure for both their group attributes and personal weaknesses, it might be expected that personal weaknesses would attract lower marks than group attributes. Overall, however, the *Primary* claims indicate that students and bands were very discriminatory in their assessments of personal weakness attribute criteria.

Secondary claims, being next in strength to *Primary*, arose from the grounded theory analysis of statistical tests performed on the data. There was a moderate correlation (i.e. 88 out of 104 assessments achieved a coefficient measurement of 0.75 or higher) between students' self-assessments of their personal weakness with those awarded by their band members. Similarly, there was a moderate correlation overall (i.e. 29 out of 34 assessments) between students' self-assessments of their personal attributes with those awarded by their band members. These findings, expanding on previous discussions involving selection of personal attributes, supports the view that there was a strong relationship between students' selection and self-assessment of their personal weaknesses with those selected and assessed by their peers. Indeed, the regularity with which bands identified personal weakness attributes that were similar to those hitherto self-selected by their individual members was, at times, remarkable.

Both group attributes and personal weaknesses exhibited small dispersals of marks overall (as measured by standard deviation). This suggests that there was strong marking agreement among students when peer assessing both types of attributes. Ultimately, however, group attributes exhibited the smallest standard deviations and this supports the assertion made in Chapter Four and elsewhere, that group attributes engendered the feeling of togetherness and shared values within bands.

Finally, *Tertiary* claims (arising from correlation analysis involving *Secondary* and *Primary data)* are considered. There were some interesting claims resulting from correlation analysis of students' self-assessments with those of their bands. This reveals a relationship between students who had under-rated themselves in their self-assessments and being subsequently awarded a high mark from their band

(accompanied by a strong band agreement over the mark). Conversely, students who highly over-rated themselves in their self-assessments (62 instances of this occurring) tended to receive a lower mark from their band (accompanied by a somewhat less strong agreement over the mark). This suggests a level of honesty being apparent among band members.

A similar relationship was observed for free riders and shooting stars also. Free riders typically over-rated themselves in their self-assessments when compared with how their bands assessed them. The reverse was found for shooting stars: they typically under-rated themselves in their self-assessments compared with those awarded by their bands.

In the previous discussion involving friendship marking, suspicions about the occurrence of pre-meditated marking agreements between particular students were raised. *Tertiary* data analysis also suggests incidents of pre-meditated marking agreements. Band members were given the option, in certain rehearsing and performing cycles (described in Chapter Four), to determine the individual intra-peer marks for each member in collaboration. On those occasions, the analysis indicates a relationship between such collaborative marking settings and an equal apportioning of band members' marks (for example, a four-piece band awarding each of its members a mark of five across all personal attributes). There is something to admire about a pop band's ethics that causes its members to agree to mark each identically, despite the very unlikely occurrence of equal contribution to the rehearsals. This phenomenon is common in all subject areas, however, with reports of 'mark-fixing cartels', for example, described in Heywood (2000, p.377). Unfortunately, as has already been remarked, not discriminating between individuals' personal attributes criteria through collaborating in this way, renders much of the care, preparation and time that were given to their formulation, redundant.

Correlation analysis involving bands' performance marks, often interpeer assessed by year three peer panels, produced some interesting findings. Firstly, it was discovered that a high band performance mark often correlated with a high consistency of intra-peer marking agreement. Secondly, the opposite of this was found: a low consistency of intra-peer marking agreement often correlated with a low band performance mark. In other words, it is proposed that band members who agree with one another about their personal attribute marking are also, as a band, likely to obtain a high mark for their

performances (and the opposite of this is similarly proposed). Perhaps it too simplistic to suggest that whenever like-minded and agreeable pop musicians are placed together in a band, they will always perform exceptionally. There are plenty of accounts documented in the public domain of bands with a history of creative differences and tensions in rehearsal, yet are recognised also as making an outstanding contribution to pop music (*The Kinks, The Who* and *Oasis* are among the standard examples). In the context of this investigation, however, there was a relationship between individuals who agreed in their intra-peer marking of each other, in rehearsal, and giving a successful on-stage performance.

A number of other relationships, involving a comparison of standard deviations with other measures within the tertiary data, were examined. The analyses of these relationships indicate several findings that together lead to two propositions. Firstly, there is a relationship between strong marking agreements of personal attributes among band members and receiving high final individual marks (together with strong agreements in other similar measures). Secondly, the opposite of this is proposed: a weaker marking agreement of personal attributes among band members is related to low final individual marks (together with weak agreements in other measures also). Such propositional claims are similar to those arising from the *primary* and *secondary* data where, for example, *shooting stars also exhibited stronger consistency of marking agreements among band members than free riders.*

The remaining analysis involved an examination of students' individual normalised difference coefficients. There was a three-way correlation relationship: an individual normalised differential coefficient that was small (meaning that the proximity of a student's self-assessment to that of the band members' assessment was of some distance) related to self-assessments that were over-rated, and also, to low final individual marks. In other words, this analysis supports the general findings that were explained in the previous paragraph.

Finally, the analysis of the tertiary data supports the earlier claim in which intra-peer marking in a collaborating setting, tends to result in bands distributing the marks equally among their members, therefore disregarding individual differences. Interestingly, the analysis also suggests that if a band contains a would-be *free rider*,

then that unfortunate individual was often excluded from the collaborative marking cartel, receiving instead, a much lower mark.

5.2.3 Summary of the claims

There was consistent agreement among students in their marking of group attributes; these attracted higher marks also than those awarded for personal weaknesses. Students generally agreed with each other in their marking of personal weaknesses, shooting stars and, to a much lesser extent, free riders. Free riders, however, were awarded poor marks in equal measure for their group attributes and personal weaknesses.

There was a moderate correlation between students' self-assessments of their personal weaknesses and those awarded by their band members. There was a moderate correlation between students' self-assessments overall of their personal attributes and those awarded by their band members.

The more students under-estimated themselves when self-assessing, the more probable that they received a high mark from their band members (about which they strongly agreed). Conversely, students who over-estimated themselves when self-assessing tended to receive a lower mark although band members' agreements were not as strong. A similar relationship was found for free riders and shooting stars, where free riders usually over-estimated their self-assessments compare with that of their band members. The opposite was found for shooting stars: they usually underestimated themselves when self-assessing compared with that of their band members.

5.2.4 Interpreting the claims

Claims arose from students' marking, derived from the knowledge of their peers acquired during rehearsing. In this sense, the signification of such claims culminates in a kind of informational knowledge that is not just of interest to the enquiry, but also for each individual student as a part of their peer feedback. From these claims, four elements of the quantitative data can be identified: group attributes; personal weaknesses; free riders; shooting stars. Each of the elements will be considered in turn, which involves an examination of the analysis and its implications for individual students.

Students agreed most of all during their assessment of group attributes. Not only was the strength of agreement among band members extraordinary, but a strong relationship existed also, between individuals' self-assessments and those of their band. This relationship suggests that individuals' understanding of their achievements, measured by group attribute criteria, differed little from their bands' perceptions. Can students learn from this, however? Perhaps individuals when receiving their group attributes marking feedback would at least be reassured that perceptions of themselves would be largely confirmed through their bands' marking feedback. There was also a moderately strong agreement for the most part when assessing personal weakness attributes. Although these marking agreements were not as strong as for group attributes, students' self-assessments of their personal weaknesses did generally correspond with those awarded by their band. There were a number of exceptions, however, and the largest variances tended to be associated with students who could be considered free riders. Individuals' understanding of the marking feedback for their personal weaknesses might involve an examination of marks awarded for these and also a comparison of these with their self-assessments. Reflecting on these variances may help individuals to develop their self-awareness, leading to a better understanding of them. If so, reflection on this peer feedback supports individuals' learning and self-knowledge.

Most noticeable of free riders[5] was the regularity with which they over-estimated themselves in terms of self-assessing. Free riders consistently awarded themselves higher marks compared with those awarded by their bands. Several bands appeared to find it difficult to agree on free riders, in view of the frequency with which a wide spread of marks was produced. In several cases too, bands appeared to be less discriminatory between the group and individual attributes of free riders, with instances of the same mark being awarded across each of these attribute types. For free riders, there would be considerable scope for them to acquire new self-knowledge from the experience involved in receiving poor marks and in reflecting about their over-estimation when self-assessing.

Indeed, the quantitative data suggests that of the 16 free riders identified at the outset of their respective cohorts during the period 2001-2005, only 11 remained as such by the time of their final intra-peer assessment. Five former free riders (EN, MS, GD, JG, TJ)

[5] As defined in Section 1.3.6.

received improved intra-peer marks indicating that they no longer belonged to that category, as a possible consequence of learning from this experience.

The strongest consistency of marking agreement overall was that which involved the assessments of shooting stars (see section 1.3.6) especially their group attributes. Shooting stars tended to under-estimate themselves in terms of their self-assessments when compared with those awarded to them by their band. Although in different circumstances from that of free riders, shooting stars themselves would acquire new self-knowledge relating to their achievements, involving, for example, the development of confidence in their abilities.

A summary of the quantitative data and its interpretation appears below as Figures 5.7, 5.8, 5.9, and 5.10.

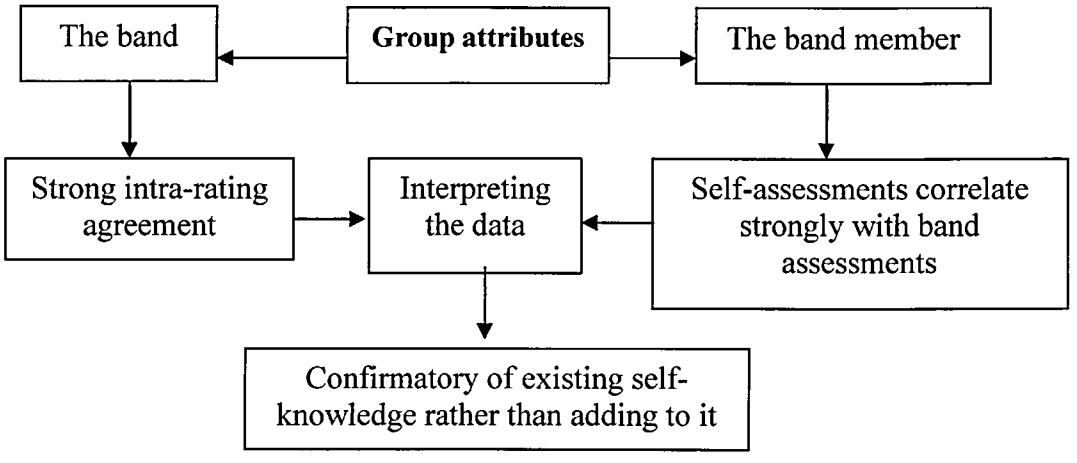


Figure 5.7 Interpretation of the quantitative data: group attributes

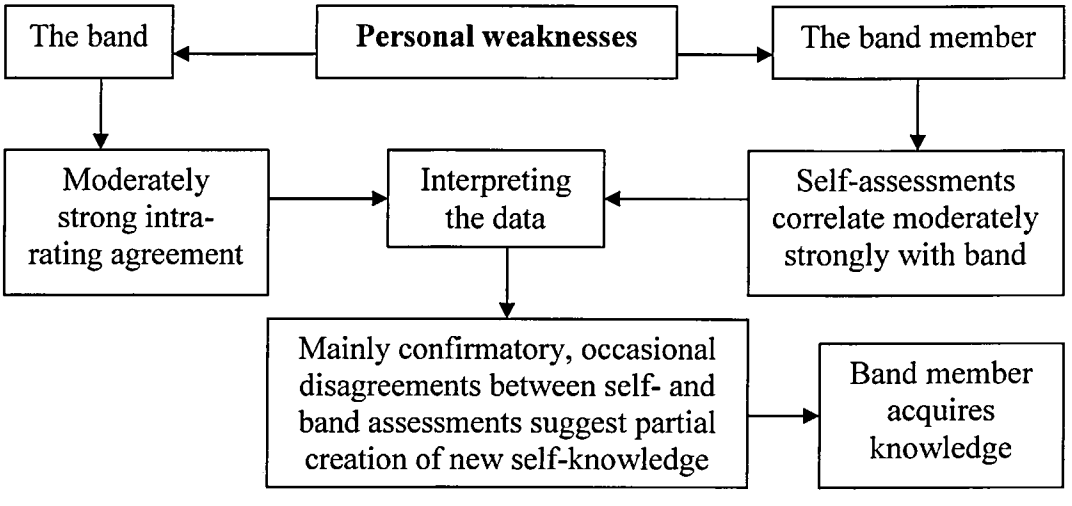


Figure 5.8 Interpretation of the quantitative data: personal weaknesses

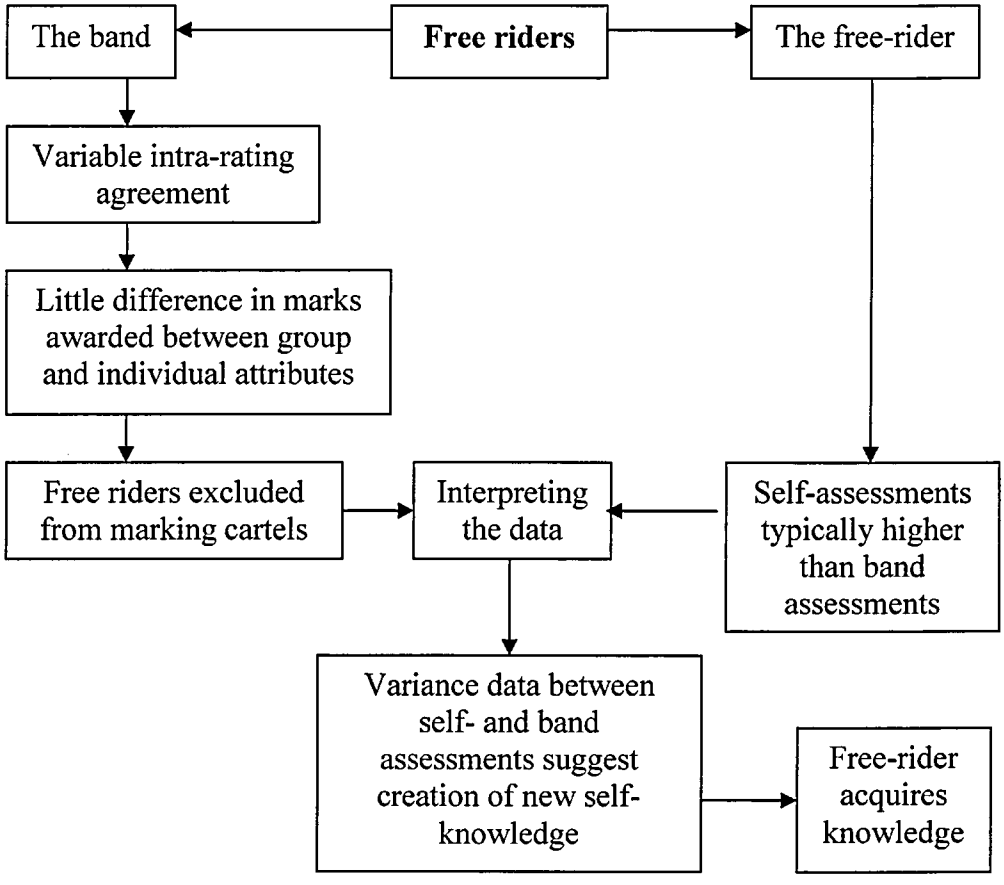


Figure 5.9 Interpretation of the quantitative data: free riders

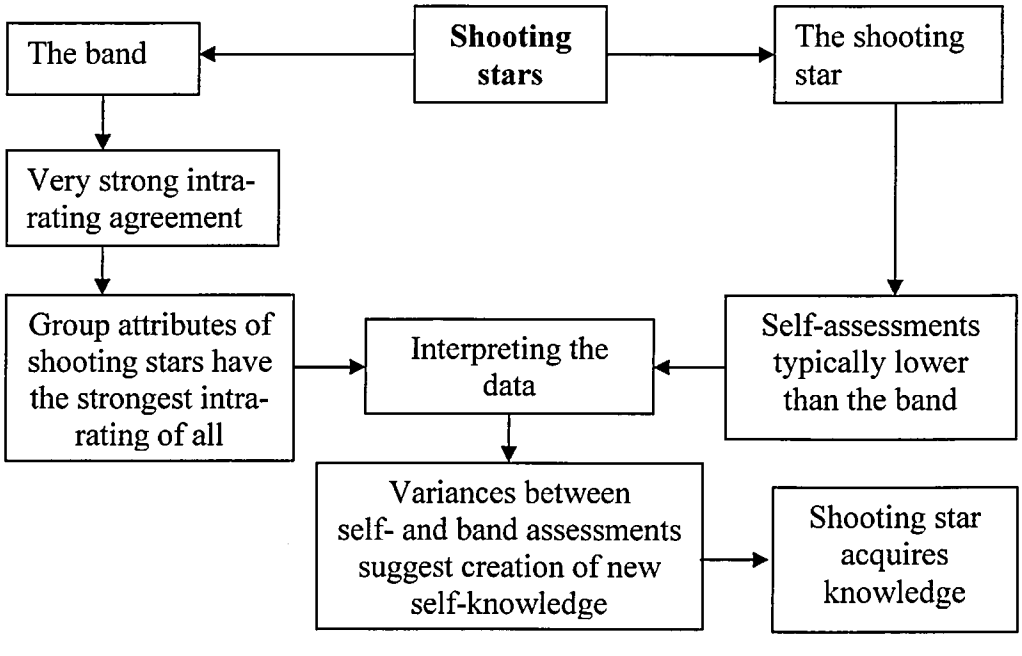


Figure 5.10 Interpretation of the quantitative data: shooting stars

5.2.5 Summary of the qualitative and quantitative data analysis

This section of the chapter concludes with a summary of the data analysis of the qualitative interview data and quantitative assessment data. The former, central to this naturalistic enquiry that is underpinned by constructivist theory, suggested that four key areas are involved in explaining interview data: self-knowledge; feedback; developing confidence; honesty. The quantitative assessment data provided a useful informational and background corpus that was detailed and extensive. Interpretation of this data analysis revealed that accessing and acquiring self-knowledge were among the key themes arising from peer assessment, including: assessment of group attributes confirms existing self- knowledge rather than adding to it; assessment of personal weaknesses can be confirmatory of existing self-knowledge, although variances between self- and band assessments suggests partial creation of new self-knowledge; self-assessments by free riders reveal they over estimate their achievements; variances between the self- and band assessments of free riders suggests the creation of new self-knowledge; shooting stars, tend to under estimate their achievements when self-assessing; variances between the self- and band assessments of shooting stars suggests the creation of new self-knowledge, which may improve their confidence. Grounded theory procedures were adapted in order to analyse both qualitative interview data and quantitative assessment data.

5.3 Relationships between the qualitative and quantitative data analysis and the interpretation of these

This chapter would not be complete without an attempt to examine and interpret potential relationships and synergies between the two data types and their analyses. A number of points of intersection arising from their relationship are suggested. These are examined in this section as follows:

5.3.1 Reflection on feedback;
5.3.2 Process features;
5.3.3 Peer marking;
5.3.4 Free riders;
5.3.5 Honesty;
5.3.6 Personal attributes and unexplained data.

5.3.1 Reflection on feedback

The analysis of the qualitative data indicates that students acquire important self-knowledge through experiencing two intra-peer assessment activities: self-identification of personal weaknesses and band identification of one's personal weaknesses. Clearly, individuals who can reflect on the variance data and what such variances between their *personal weakness* self-assessments and the intra-peer assessments given by their bands might mean, learn important knowledge about themselves. As there was widespread agreement among interviewees about the fairness of their bands' assessments, every opportunity should be made available for students to be able to reflect on the variance data arising from each of their personal attributes. If so, such reflection may allow individuals to be able to access a particular kind of knowledge about themselves that is difficult to obtain through conventional modes of assessment or discourse.

5.3.2 Process features

Interviewees' responses involving the process features line of enquiry (discussed in Chapter Four) were analysed. Three process features were involved: conducting marking in private or collaborative settings; the timing of the intra-peer assessment; the status of the marks. Firstly, the interview analysis indicated a strong preference for

private marking (explained, for example, in *The Focus on Honesty*), as students considered that marking in private encouraged greater honesty of assessment feedback. Secondly, many interviewees favoured intra-peer assessment that was conducted prior to the performance. Although the majority expressed a view that the timing of it would not influence their marking, no interviewee advocated conducting intra-peer assessment post-performance. Finally, the analysis indicated interviewees' preference that marks should count towards end of year one grades, for a variety of reasons (rehearsing is a major part of the module, they would take the process more seriously, for example), but perhaps not to contribute to final degree classifications.

5.3.3 Peer marking

Of course, not all interviewees agreed with the marks that were awarded by their peers. The data suggests that the major disagreements about the marks that were awarded involved free riders (indicated by interview comments and intra-peer marking) whose frequent and unexplained absence or lack of commitment in other ways was a source of resentment by other band members. Interestingly, shooting stars were very rarely subjects of animosity, arising out of jealousy; rather, it was the opposite. The qualities of shooting stars as contributors to rehearsing, often being regarded as natural bandleaders, were admired. This is indicated by the frequency with which bands selected personal attribute criteria for their members such as *leadership skills*, *confidence* and *organisational abilities* – qualities characteristic of shooting stars, as did individuals themselves when determining their self-chosen personal weaknesses.

The resentments and the interpersonal tensions that appeared in the rehearsals, reported in the interviews with free-riders (e.g. EN, WK) and recalled by other interviewees (e.g. EJ, NL) who encountered free-riders in their bands, anticipated difficulties ahead for peer marking and for free-riders' responses to their receiving of those marks. Indeed, bands were less assured in their marking agreements (as measured by standard deviation) when assessing would-be free-riders than either shooting stars or all remaining students.

Although pre-meditated intra-peer marking agreements would not be easily prevented, I did think that it might be detectable. I often suspected that some kind of pre-meditated friendship marking agreement took place when observing extraordinarily similar or

implausibly identical intra-peer marks between known friends. Another reason for deciding to discontinue with collaborative marking (as well as discomfort expressed by a number of students in their interviews, described previously) was that it might facilitate friendship marking. There were a number of instances, for example, of band members uniformly awarding each other straight threes or fives across all personal attribute criteria, therefore rendering the particular categories, in which so much time was devoted to their identification, irrelevant. However, I only recalled one student complaining about suspected friendship marking at the time, other than those who raised it in the interviews. Similarly, no student reported that they considered themselves disadvantaged, in terms of the marking, as a consequence of their not belonging the circle of friends within their band. It had been explained to students as a part of peer assessment rationale, that I (or other tutors) would moderate marks where required (for example, a suspicion of unfair marking) and they had could report their suspicions to myself or other tutors. At least one interviewee (EJ) suggested, "unfair marking along personal dislike should be discussed"; indeed, such a discussion might contribute towards promoting the principle of transparency and honesty that is espoused in the pre-requisite stage of the intra-peer assessment process model.

5.3.4 Free riders

Although no interviewees questioned the authority of their band to select and to intra-peer assess their personal weakness attributes, being able to accept the decisions of the band was for some free-riders[6], as described in their interviews, a difficult learning experience. Such responses illustrate some of the tensions that may occur in employing band-determined personal weaknesses and the difficulties that may be faced when engaging in this particular approach to peer assessment. Learning through such engagement, in which students are able to access and acquire critical self-knowledge through reflecting on feedback, for example, can be regarded as activities that typify social constructivist pedagogy.

In continuing the analysis of free riders, the intra-peer assessment data suggest that sixteen individuals, out of the total population of 170, could be defined as being a free rider at some point in at least one of their bands. Two of the students who were interviewed were regularly free riders (EN and WK) and a third interviewee (TP) fell

[6] As defined in Section 1.3.6.

into the free-rider category in one of the bands in which s/he appeared. These students make a number of points in their interviews relating to how they perceived the assessment data. EN, for example, appears to suggest that the material his/her band was asked to perform was "not suited" to his/her performance style, asserting that this was a major reason why s/he received poor marks. WK made a similar point, in stating that disagreements over choices of songs led to his/her attributes being "marked down". WK, in having *Creative input* as a personal weakness, considered that developing this attribute was difficult as the band wished to faithfully copy the commercial recording of the artist, therefore restricting opportunities for demonstrating improvements in his/her musical creativity through, for example, arranging the song.

These examples are illustrative, perhaps, of decisions by the band during rehearsals, involving changes that have consequences for individuals in being able to demonstrate particular attributes that had already been selected. This is an important point that has implications that go beyond the choice of criteria. It raises questions also about whether such criteria might need to evolve during the rehearsals in order to reflect a band's changing rehearsing or artistic direction. However, do such changes impinge upon the raison d'etre for the original choice of personal attribute criteria? These questions are explored further in Chapter Seven.

Apart from the implications for free riders' motivation, these interview comments may explain some of the wide variances that were found between their self-assessments and the intra-peer assessments that were made by their bands. Were these variances that were produced a result of students' differing perceptions of personal weakness attributes as a consequence of changing rehearsing contexts and experiences? In combining both interview and assessment data analyses, the following interpretation may be suggested: free riders were self-assessing their personal attribute criteria in terms of what they believed they had agreed upon prior to the rehearsal, but their band were assessing the meanings of these qualities in an experiential context of rehearsing that was developing and evolving. Such an interpretation has significance also, for constructivist pedagogy; this is discussed further in Chapters Six and Seven.

5.3.5 Honesty

Honest is a key theme emerging from both the analyses of both data types: it is a core category with thirteen interviewees preferring honest marking in private, marks counting towards a summative assessment. This preference is illuminated by the quantitative data analysis also. For example, an examination of the variance data between the self-assessments and band assessments of each individual can be regarded as one measure that is indicative of the extent of honest marking. If, for example, variances had been large or the intra-rating agreements had been weak among extensive numbers of students, then this would cast some doubt on the honesty and integrity of students' peer marking. The variance and intra-rating data suggests that this was not the case.

5.3.6 Personal attributes and unexplained data

This section concludes with a discussion of the personal attributes that individuals frequently formulated and their relationship with findings that are unexplained. One example involves the quantitative analysis of group attributes data, where it indicates that intra-rating agreements of peer assessments are at their strongest. Related to this are comments made by some interviewees, in which they appeared not to differentiate between group attributes and other attribute types when marking. Although these interviewees did not articulate the distinctive character of their group attributes, differences in the choice of names that they ascribed to these were apparent, compared with other types of personal attributes. For example, the most frequently employed group attributes included: *Team working; Commitment; Attendance; Organisation.* These qualities did not appear among the most frequently selected personal weakness attributes. In contrast, the most frequently selected names of personal weakness attributes included: *Confidence; Contributing to ideas; Enthusiasm.* It may be suggested that the very nature of group attributes (meaning group-agreed attributes, as previously explained), being predominantly characteristic of group working or team building behaviours or traits, had a causal effect on the intra-rating strength of marking agreements. An alternative explanation might involve the experiential context in which the band, as an entity, decided which mutually agreeable attributes were important for them. If so, strongly held group agreements about what they considered were important for them, were demonstrated at the point of assessment by strong marking agreements

about their intra-peer group attributes. It is an interpretation that might, as such, resonate with social constructivism.

To summarise then, this chapter attempts to make sense of the qualitative and quantitative data corpus. The chapter has presented an analysis of the responses and views of the students and their bands to peer assessment. It offers interpretations of the data analyses arising from the action research that has significance for the enquiry and beyond. The substantive categories arising from the interview data were as follows: self-knowledge; honesty; feedback; confidence. Propositional claims arising from the assessment data indicated the following categories: group attributes; personal weaknesses; free riders and shooting stars. Grounded theory principles were adapted to analyse both data types. This involved the development of a number of procedures considered appropriate for the analysis of the quantitative of peer assessment data.

Chapter Six: 'Knowing yourself through others': a new process model for intra-peer assessment

6.1 Presenting a new process model of intra-peer assessment

6.2 Stage one

6.3 Stage two

6.4 Stage three

6.5 Stage four

6.6 Using personal attributes as assessment criteria: developing awareness and supporting learning

6.7 Moral dimensions of the process model

6.8 The process model and band determined personal weaknesses

6.9 Strengths and weaknesses of the model

6: 'Knowing yourself through others': a new process model for intra-peer assessment

Chapters Four and Five have presented an analysis of the developing rehearsing and performing cycles and the qualitative and quantitative data that these produced. This chapter proposes a new process model for peer assessment that synthesises the major findings of the action research in this enquiry.

The chapter begins with an introduction and overview of the stages that comprise the model. A discussion of each of the four stages in turn, accompanied by diagrammatic illustrations, then follows. It continues with an examination of the fundamentals underpinning the model, including the use of personal attributes as intra-peer assessment criteria and its moral dimension of trust, honesty and fairness. A key activity, and one that gives the model a particular distinctiveness, occurs in stage three whereby bands decide for each of their members their personal weakness on which they will be intra-peer assessed. This activity is central to the model and a detailed discussion appears in Section 6.8. The chapter concludes with an examination of the strengths and weaknesses of the model. The chapter is organised into the following sections.

6.1 Presenting a new process model of intra-peer assessment
6.2 Stage one
6.3 Stage two
6.4 Stage three
6.5 Stage four
6.6 Using personal attributes as assessment criteria: developing awareness and supporting learning
6.7 Moral dimensions of the process model
6.8 The process model and band determined personal weaknesses
6.9 Strengths and weaknesses of the model

6.1 Presenting a new process model of intra-peer assessment

Over the duration of the five-year long action research period, a new process model of intrapeer assessment has developed. The evolution of this model has been a gradual process that has arisen out of the rehearsing and performing cycles and from the analysis. The fundamentals of the model, illustrated in Figure 6.1 (p.204), are a sequence of activities, comprising four stages. The rehearsing and performing context and the role of the tutor are central to this. The model encourages students and the tutor to experience, share and develop awareness and knowledge about themselves and about their band members in rehearsing and performing through intra-peer assessment activities. Such experiences and activities may be considered as belonging to social constructivist epistemology, and the importance of constructivism in the context of this model is discussed in Chapter Seven.

Integral to the model is the tutor's role in facilitating experiences in which student learning and knowledge creation may develop and progress. This includes providing opportunities for students to: develop an awareness of their own *personal weaknesses*; reveal these *personal weaknesses* to their band members; agree to these being used as intra-peer assessment criteria; decide as a band the *personal weakness* criteria for each band member; intra-peer assess other band members' *personal weaknesses;* reflect upon and respond to intra-peer assessment feedback.

The model, then, offers a sequence of graduated stages of personal attribute usage, which create experiences that support students' learning about themselves and about others in intra-peer assessment activities. It represents a significant step forward in our understanding of peer assessment. The model forms a pedagogy, which leads the individual from a limited awareness of one's own personal attributes, limited self-assessment and limited feedback, to experiences that develop each of these. Students move from a position of revealing a little about their own personal attributes in group rehearsing to one in which they agree to others (their bands) deciding the personal weaknesses that ought to be used as assessment criteria for them. It is a journey that requires careful preparation and one that defines the distinctiveness of this research in which preparation, care, sensitivity and responsiveness is required. Values of a moral kind are also developed, including developing trust and honesty through engaging in various activities that involve the intra-peer assessment of personal

attributes. The process model also fosters students' confidence in group-work activities and, through intra-peer assessment activities, develops their maturity. These aspects are discussed later in the chapter.

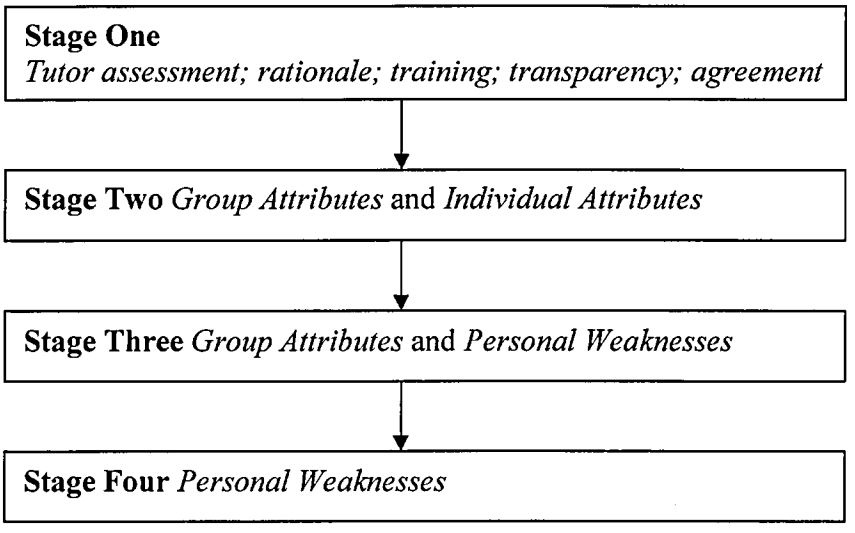


Figure 6.1 Stages of the intra-peer assessment process model

6.2 Stage One

Peer assessment can be stressful for students (Pope, 2001, 2005) and this may be exacerbated by the absence of careful preparation for engaging in such activities. This section explains the activities at stage one that were established in this study that are proposed as a preparation for stage two. These comprise: involvement in an initial tutor assessed group rehearsing and performance assignment; explanation of the rationale for intra-peer assessment; the provision of peer assessment training.

Students should be involved in a conventional tutor assessed group rehearsing and performance assignment prior to experiencing peer assessment activities. An initial tutor-assessed assignment enables students to comprehend performance criteria and performance standards, as well as experiencing group rehearsing with unfamiliar band members without the added pressure of peer assessment. For example, the first assignment for each of the five cohorts in this study, completed by the middle of the autumn term, did not employ peer

assessment. For this first assignment tutors simply awarded marks to bands' performances as a whole, rather than to individuals. Introducing peer assessment in the second assignment of the course (the *Christmas Party)* also allowed myself, as the tutor, an opportunity to acquire some experiential knowledge of the students and their bands prior to commencing intra-peer assessment. Obtaining experiential knowledge of the students and their bands, a key element of the model, enables the tutor to form an opinion concerning, for example, students' readiness or otherwise, to progress to the next stage. This is discussed in Section 6.3.

At least two or three hours are usually required to explain the rationale for intra-peer assessment and to provide peer assessment training involving the use of personal attributes. Presenting the rationale of the intra-peer assessment system involves explaining its potential learning benefits, reassuring students about the transparency of the process, outlining principles of agreement, describing the concept of using personal attributes as assessment criteria and safeguards such as tutor moderation.

Students should additionally be provided with one or more training activities involving the identification and formulation of personal attributes related to group rehearsing. A fuller description of the training procedures used in this study can be found in Sections 3.3.1.1 and 3.3.1.2.

Although the presentational activity of the rationale, at this first stage, is a pre-requisite for entering stage two, the activity also supports each of the successive stages in the model; this is explained later in the chapter. It is usually necessary to re-visit the rationale, emphasising once more the potential benefits of using personal attributes as assessment criteria, demonstrating transparency, and establishing the principle of agreement. These, together with the necessity for the tutor to acquire some experiential knowledge of the students, are key elements in the pedagogical underpinning of this process model.

6.3 Stage Two

At the centre of the model lies a process for the identification of particular personal attributes, thought to be important for individuals as they rehearse in their band. A considerable part of the research explored the use of students' personal attributes as criteria for peer assessment. This involved developing appropriate methods for formulating group attributes and individual attributes in the context of band rehearsing and performing. These activities are introduced to students during the second stage of the model, illustrated in Figure. 6.2.

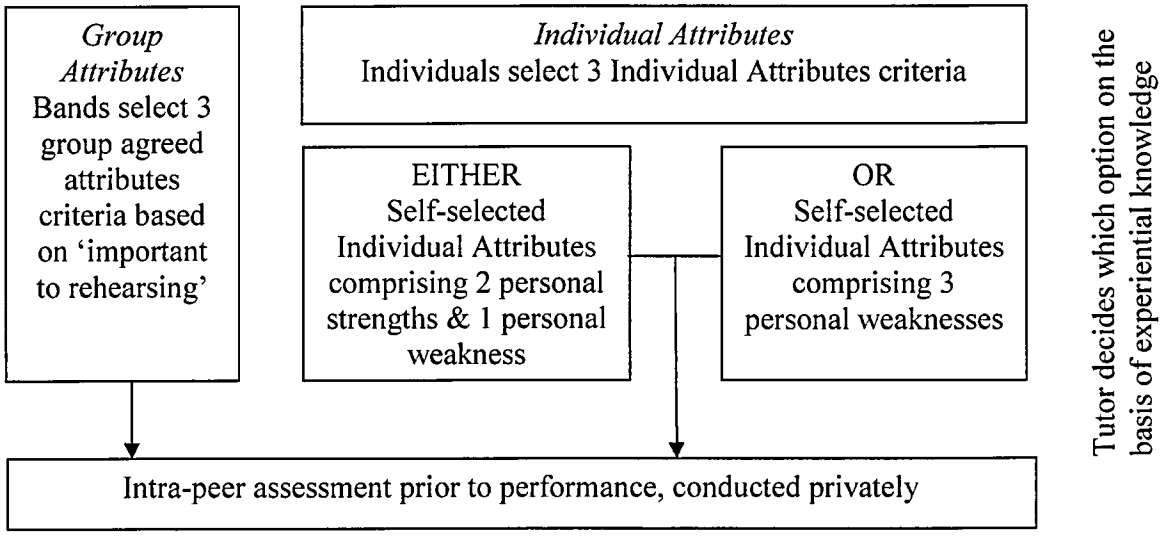


Figure. 6.2 Stage Two of the intra-peer assessment process model

At stage two students develop awareness of themselves and of others in rehearsing through establishing rehearsal-related *Group Attributes (i.e.* group-agreed attributes*)* and *Individual Attributes* (i.e. particular to each band member) to be used as assessment criteria.

Using peer assessment criteria based on students' individual attributes involves a risk of intrusion into areas of a person's character that might be regarded as private or personal. It is important, therefore, that students' agreement be sought, at the outset of the process. Tutors may wish to revisit the rationale, at this point, explaining to their students the benefits of using personal attributes as assessment criteria in terms of enabling self-improvement in rehearsals.

Group Attributes, with their focus on what is important for the band, are less intrusive for individuals because they are formulated and agreed by the group as an entity. As such, group attributes are suitable as the starting point, at stage two, for the activity of identifying personal attributes. As well as introducing students to the process of identifying and establishing these, group attributes, being founded upon mutually agreed principles, symbolise shared band values, which can help to develop trust. Developing trust, among other moral dimensions, is important for the model; this is discussed in the following sections and also in Chapter Seven.

Individual Attributes, involving a focus on specific rehearsing qualities of each band member, are introduced at stage two in order to develop self-awareness at the outset. Students formulate these by themselves, on the basis of identifying their own personal strengths and personal weakness in rehearsing. Stage two offers a choice for the tutor in deciding which kinds of individual attributes are most appropriate for the class as a whole: self-selected individual attributes comprising two personal strengths and one personal weakness, or, self-selected individual attributes comprising three personal weaknesses. Rather than using the term 'personal weaknesses', tutors may instead prefer to express these as 'qualities needing improvement or development' in their exchanges with students. This expression, with its similar emphasis on learning, may offset the negativity that may be implicit when using the term *personal weaknesses.*

The first choice, utilising two personal strengths and one personal weakness attribute, was employed during the rehearsing and performing cycles of 2003. This option offers a more cautious and graduated sequence of steps for supporting students in disclosing their personal weaknesses that is less threatening in terms of personal intrusion.

Alternatively, the second choice involves the individual self-selecting three personal weakness attributes at the outset, by-passing the above step that combines strengths and weaknesses. Self-selecting three personal weaknesses at stage one, progresses students more speedily through the stages of the model. It is for the tutor to decide, on the basis of her/his knowledge of the students' 'readiness' or maturity, which of these two choices is the most appropriate.

A number of process features of peer assessment have been extensively explored during the action research cycles. These features or operational mechanics have attracted relatively little attention in peer assessment studies (as surveyed in Chapter One) and the findings of this research adds significantly to our understanding of the effect of such processes on peer assessment systems. These findings are incorporated into stage two and are embedded throughout the model. Firstly, intra-peer assessment is at its most effective, in terms of encouraging honesty of marking, when conducted after the final rehearsal but before the performance. Secondly, and in order to promote further honesty, intra-peer assessment is best conducted with students assessing in private. Thirdly, tutors should decide, on the basis of their experiential knowledge of the students, whether or not to use the marks for formative or summative grading purposes. The data analysis in Chapter Five indicated also that 'marks should count' towards the end of year grades, 'but not towards degree classifications' and summative marking fostered a more 'serious' student engagement with intra-peer assessment. It is therefore proposed that the model incorporates summative marking, in order to provide individuals with a measurement of their attainment, that contribute to end of year grades, but not to degree classifications.

An important activity which is encountered during stage two is that in which individuals are encouraged by the tutor to identify and formulate their particular personal attributes and consider agreeing to these being used as intra-peer assessment criteria. Revisiting the rationale, the transparency of assessment, the principle of agreement and moral judgements about levels of honesty and trust are significant in students enjoying a rewarding learning experience at stage one.

6.4 Stage Three

When sufficient student-tutor and student-student trust and confidence in the process have developed and the tutor has also acquired sufficient experiential knowledge of the bands, then progression to stage three of the model, as illustrated in Figure 6.3 (p.209), can be contemplated.

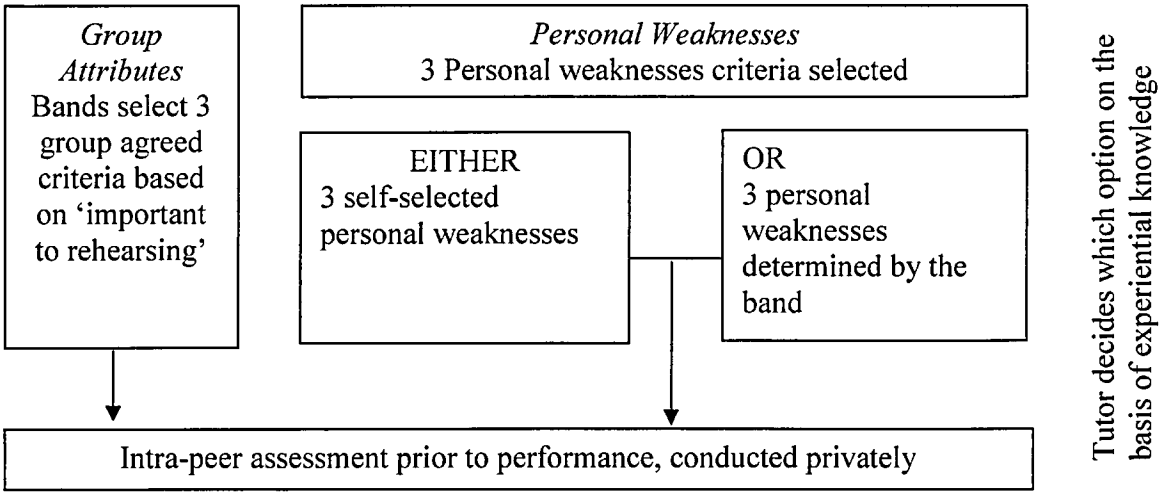


Figure 6.3 Stage Three of the intra-peer assessment process model

Stage three occupies a central place in the model; it is here that students are introduced to a key activity of the model in which they assign responsibility for the selection of their personal weakness attributes to the band with which they have rehearsed. This is a critical moment and one that depends upon the tutor's experiential knowledge of the band members, in deciding whether sufficient trust and honesty has developed that would encourage students to agree to their participation.

It is natural for students' unease and doubt to surface at this point. Difficulties may be faced in both extracting and receiving personal information of this sort: for bands in deciding the personal weakness attributes of others and for individuals also at the receiving end. Although preparation for intra-peer assessment in this study (i.e. stage one) occurred in advance of rehearsing and performing, the rationale was regularly revisited, re-emphasising transparency and seeking agreements, especially during this critical stage. It was also appropriate on occasions to explain, once again, the potential learning benefits for students through the assessment feedback that they would receive.

Stage three (as in stage two) allows the option for tutors to defer progression to band determined personal weaknesses, where tutors are unsure of the 'readiness' of their students. If so, there is the option to repeat the activity involving self-selected personal weaknesses, in order to allow confidence, trust and honesty to develop further, along with the tutors' experiential knowledge of their students.

Formulating the three group attributes continues as before and, in addition to maintaining continuity with stage two, this activity also helps to build further trust within the bands. Intra-peer assessment is conducted in private and in advance of the performances, as before.

6.5 Stage Four

In continuing the increasing focus on the individual learner that was introduced at stage three, stage four also changes the balance of group attributes and band-determined personal weaknesses, as illustrated below in Figure 6.4. For stage four, the assessment is weighted in favour of band-determined personal weaknesses in order to raise their importance for individuals in developing awareness and further motivation. At stage four, the tutor is presented with two choices: continue using group attributes, in the form of a single tutor-imposed 'contribution to rehearsing' criterion, or, to discontinue with the use of group attributes. The consequences of either choice is a reduction in the weighting for group attributes from either 50% (in stage two) to 25% if deciding on the first option, or from 50% (in stage two) to 0% should the second option be chosen. Using a 'contribution to rehearsing' criterion enables the other band members to provide a general, overall intrapeer assessment feedback of the individual's commitment and support of the group, alongside their personal weaknesses attributes.

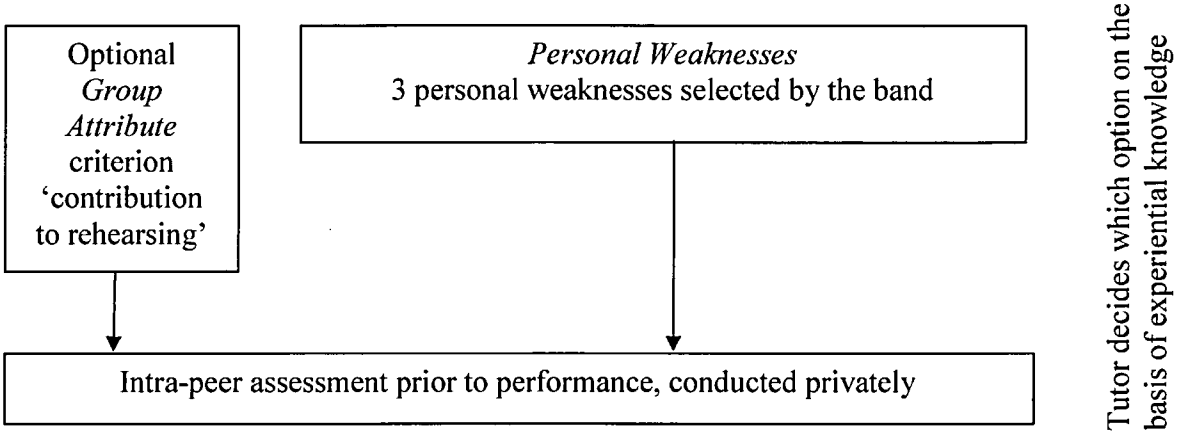


Figure 6.4 Stage Four of the intra-peer assessment process model

Tutors should decide whether their students would respond positively to intra-peer assessment based solely upon their band-determined personal weaknesses, or to continue with using a limited group attributes criterion based on 'contribution to the group'. Dispensing with group attributes in favour of focusing intra-peer assessment exclusively on individuals' band determined personal weaknesses increases students' awareness, importance and motivation to improve these. This refocus, however, is at the expense of individuals receiving feedback from their band that reflects other aspects of their contribution to rehearsals. Selecting this approach, based purely on band-determined personal weaknesses and discounting any other contribution to rehearsing, is decidedly the more risky option, although it seeks to obtain the greatest learning benefits for the individual. The benefits and risks involved in each approach are finely balanced: it is for the tutor, on the basis of his/her experiential knowledge of the class and in consultation with the students, to decide which option.

The peer marking activity, as in stages two and three, continues to be conducted in private, preceding the performances. Stage four is the culmination of the student's journey through which they learn about themselves and develop their personal attributes. It is a carefully structured sequence of activities that are designed to develop important areas of learning over a realistic period of time, not a 'quick fix' or 'one-off' student experience. The four stages comprising the intra-peer assessment model are illustrated together in Figure 6.5 (p.212). This model expands our understanding of the pedagogy of popular music group work involving rehearsing and performance. It offers a distinctive approach to intra-peer assessment that supports learning and progression through developing an awareness of personal attributes and bringing about improvements in key areas identified by honest and trusted peers.

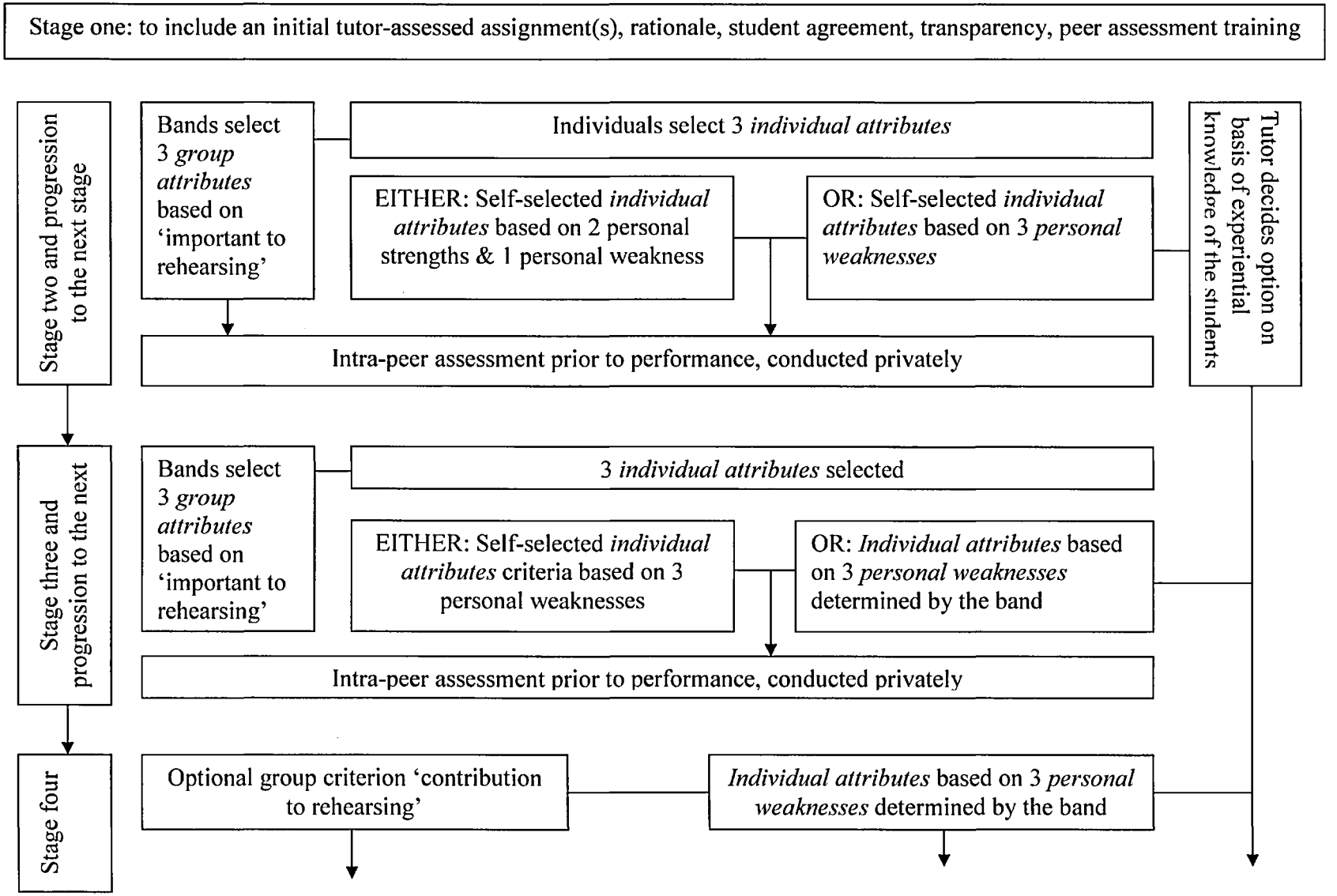


Figure 6.5 Overview of the four stages of the intra-peer assessment process model

6.6 Using personal attributes as assessment criteria and the process model

The research indicates that students' self-awareness and awareness of each other's personal attributes seems to have increased as a consequence of their involvement in intra-peer assessment and, moreover, self-knowledge became a substantive category arising from the qualitative data analysis. The process model develops awareness, in part, through engaging students in activities related to formulating personal attribute criteria. This includes self-determined intra-peer assessment criteria and, through participating in band discussions, deciding personal attribute criteria for others. Further opportunities for developing self-awareness are provided by the model through feedback where students can compare their self-selected personal weakness attributes with those that were determined for them by their band.

Identifying how, which and why particular personal attributes may show more potential than others in bringing about improvements to learning, in a group context, when used as assessment criteria is problematic. Assembling a checklist of particular personal attribute criteria, for example, *Contributing to ideas* is somewhat simplistic. Devising checklists, from the data analysis of this research may lead to making assumptions that a personal attribute such as *Contributing to ideas* can be used for supporting learners generally. This is problematic because it raises questions about the transferability of personal attribute criteria across different learning contexts (for example, in other bands) and for different individuals. Making assumptions about the transferability of key skills is a highly contentious area (Neath, 1998) and, as explained in Section 2.2, the use of personal attributes in this enquiry is highly situational and experiential. The personal attributes in this study were conceived in the context of first year undergraduate popular music degree students within the shared experiences of their particular band rehearsals. For these reasons, making assumptions about the transferability or universality of personal attributes criteria to other contexts is problematic.

It is also suggested, in Section 5.2.2, that a 'tacit' knowledge or common understanding of the meanings of personal attributes was created, rendering as unnecessary a need for these to be always defined discursively in the class. The creation of a common understanding of what is meant by, for example, *confidence*, *creative input* or *tolerance*, arises from the

shared experiential context of the learners. As all the learners in each cohort were engaged with the same set of contexts and experiences during their rehearsing, it is from this sharing that a common understanding of the meanings of personal attributes was developed. This view, founded on constructivist principles of created knowledge, does not support assumptions about the transferability or universality of personal attributes, in the sense of the previous paragraph. It also raises the question of whether symbolic interactionist theories might help to explain the meanings that students attach to their personal attributes. This is explored further in Chapter Seven.

The previous chapter presented a discussion about the variety of personal attributes that arose from the analyses. This concluding section identifies and suggests possible relationships between the personal attribute analyses arising from Section 5.3, and the intra-peer process model. Figure 6.6 illustrates possible relationships between the personal attribute criteria that were the most frequently formulated by students and the model.

Quantitative data analysis: Most frequently formulated personal attributes criteria		Possible relationship to the model
Group Attributes	*Attendance; Team working; Enthusiasm; Commitment; Responsibility for learning parts; Organisation; Punctuality*	*Developing trust within the group; Stages two and three;*
Personal weaknesses	*Confidence; Contributing to ideas; Punctuality; Enthusiasm.*	*Stages two, three and four; Developing learning*
Personal strengths	*Responsibility for own part; Punctuality; Organised; Confidence*	*Stage two; Developing confidence in oneself;*
Qualitative data analysis: Personal attributes arising from the interviews		
Self-responsibility	*Commitment; Responsibility; Reliability; Punctuality; Attendance.*	*Knowledge of self;*
Interpersonal relationships	*Tolerance towards others; Personal skills; Flexibility; Team member; Communication with team.*	*Knowledge of self/others;*
Creativity	*Trying new skills/ideas; Inventive; Creative input; Contribution to ideas*	*Motivation;*
Confidence	*Confidence; More verbal input; Leadership skills/taking charge.*	*Confidence;*
Awareness	*Listening to others; Willingness to help others; Focus/concentration*	*Awareness;*

Figure 6.6 Personal attributes arising from the qualitative and quantitative data analyses and their possible relationship to the intra-peer assessment process model

6.7 Moral dimensions of the process model

This section discusses the moral dimensions of the process model, which are so distinctive involving, in particular, those of trust, honesty and fairness. Implicit within each stage of the model is the building of trust between individuals and their bands. For example, stage two involves individuals placing trust in their band to consider their self-selected personal attributes that they have disclosed to them; stages three and four similarly invite individuals to consider trusting the moral imperatives that are implicit when band members decide appropriate band-determined personal weaknesses. Because students work together on developing mutually agreeable group attributes for their band, this activity has importance for building trust between students. On reaching stage four, there is the option, if the tutor decides that sufficient trust has developed, to discard group attributes in order to increase the focus on improving students' personal weaknesses. Of similar importance is the building of trust between the student and the tutor[1] and, although discussed previously, it is revisited also in Chapter Seven.

Honesty, as explained in the previous chapter, emerged as a substantive category arising from the analysis[2]. The desire for honest feedback and fair marking expressed by interviewees is of relevance to a moral underpinning of the process model. Indeed, the desire for honest feedback to everyone including friends emerged as a subcategory of the axial coding. Of the thirteen interviewees who stated a preference for marking in private, six interviewees were explicit about it being necessary for encouraging honest feedback. In fact, interviewees offered a variety of explanations, discussed in Section 5.1, which supported their preference for private marking. The following is a summary of these: anxieties about being placed in a situation of having to telling the truth face to face; collaborative marking could cause tension or break-up; being uncomfortable in giving low marks in a collaborative setting which also leads to the avoidance of being too harsh; marking should not be influenced by other people; disputed marks would need justifying with a tutor.

[1] Interviewees said very little about my role with regard to trust; this is discussed further in Chapter Seven.
[2] Seven subcategories indicated a focus on honesty (1, 2, 3, 5, 6, 7, 8) including four falling into the 'strong' boundary and two belonging to the 'moderate' boundary.

The process model attempts to address these concerns through its embracing of intra-peer assessment that is conducted in private, and a rationale that makes clear the availability of tutors' moderation of disputed marks. Trust and honesty are central to the formulation of self-selected personal weaknesses. Not only is it a matter of trust that individuals offer legitimate self-selected weaknesses, but it is implicit also, when allowing their peers to assess these fairly, particularly so when marking in private. Working together as a band in developing group attributes is an activity that develops trust and honesty and one that encourages individuals to offer genuine self-selected weaknesses and marking that is fair. Table 6.1 below, is a summary of the relationships involving stages two, three and four that link individuals, bands, tutors, and moral dimensions of the process model, such as trust and honesty.

	Individuals	Bands	Tutors
Stage Two	Honest choice of self-selected personal attributes	Honesty when conducting intra-peer assessment	Tutors trusted by students to ensure fairness of peer assessment system
	Trusting the band to peer assess fairly		
Stage Three	Trusting the band to determine personal weaknesses fairly	Honesty when determining individual's personal weaknesses	
	Trust in the band to peer assess fairly	Honesty when conducting intra-peer assessment	
Stage Four	Trust sufficiently established to discard group attributes and allow focus on personal weaknesses	Honesty when determining individual's personal weaknesses	
		Honesty when conducting intra-peer assessment	Tutors decide whether sufficient trust exists to enable group attributes to be discarded

Table 6.1 Summary of the relationship between trust, honesty and the process model

Matters of trust and honesty may segue into questions involving fairness and tensions that were raised among interviewees. Some interviewees used phrases such as 'biased marking' and these were sufficiently strongly felt that it resulted in the axial coding category: 'Bias', as analysed and discussed in Chapter Five. It would, however, have been remarkable had there had been no reports of suspected unfair marking, in view of the complexities of human behaviour when encountering opportunities for indulging in questionable practices.

As such, I accepted the likelihood at the outset of the action research that there might be many who would attempt to exploit or manipulate this mode of assessment for the purpose of inflating their personal marks. Although I anticipated many allegations involving unfair marking practices, I was both relieved and surprised that reports of suspected unfair marking appeared to be unexpectedly slight. Despite receiving remarkably few grievances over the lifetime of the action research, the three complaints I received might well have a deeper extent of student concerns[3]. While some of these surfaced during the interviews, the actual level of unfair marking both to victims as well as to myself was unknown. Were there instances of unfair marking that, for example, individuals decided not to bring to my attention because of an assumption that I would be powerless to be able to seek redress? Similarly, were there victims who, believing that they were not significantly adversely affected, decided against the reporting of this? If so, maybe there were individuals who did not wish to report it out of fearing further victimisation from their band. Perhaps there were victims who did not bother to report this because they preferred to deal with it themselves. This last point recalls a remark made by EJ (appearing in Section 5.1.1) that unfair marking can lead to retaliation. Resolving these difficult issues is problematic not only for the process model but also for peer assessment generally. Tutor moderation of marks, in addition to the principle of marking conducted in private, advocated in the model, act as safeguards for individuals who suspect unfair marking practices.

Analysis presented in Chapter Five suggests that unfair marking appeared to arise out of a variety of situations, rather than from a repeated or particular set of circumstances. Personal feelings related to either animosity or friendship was a prime cause of unfairness evident in the data analysis. Of course animosity between students can develop prior to rehearsing and performing and such tensions may be exacerbated through the necessity of having to work with each other.

Strained relationships between particular band members can occur in complex ways during rehearsals, and this often involves a clash of personalities. Tensions can be inherent among members of a band during the creative process; the public domain is documented with countless examples of such tensions among well-known rock bands. In the *Performance*

[3] This may raise speculation about the tutor being regarded as a figure of power and authority and whether this had a bearing on the low number of complaints received from *Performance Management* students.

Management module, disagreements involving musical taste, choices of songs, creative direction and reluctance to perform material in unfamiliar musical genres were typical sources of tension that were described by interviewees. The process model offers scope for individuals and bands to resolve these tensions through the use of judiciously selected personal attribute criteria. Whether these are self-selected or band-determined, knowing that these personal attributes will be peer assessed can lead to individuals improving, for example their tolerance of others, and assist in resolving conflict. Indeed, the regularity with which personal attribute criteria such as *Flexibility* and *Openness to different genres* were either self-selected or determined by bands, suggests that the students had recognised, in themselves, a need to address and resolve tensions between them.

The possibility of unfairness arising from marking conducted after individuals were aware of their performance mark, was explored through interventions in the timing of intrapeer assessment and described in Chapter Four. Interview data analysis of this, indicated that students thought intra-peer assessment was more effective when conducted after the final rehearsal and before the performance; the model implements this sequence of activity[4].

The issue concerning unfair feedback or criticism that is received by students requires careful thought. Tutors may need to consider questions about maturity when preparing students for receiving feedback, particularly when introducing personal weakness attributes. Addressing these issues in the training and preparation sessions prior to rehearsing is advisable. Judging whether students would respond appropriately when receiving negative feedback is also a responsibility for tutors, when deciding their readiness to progress to the next stage in the model.

This section examined the distinctive moral dimension of the model and how, for example, honesty, trust and fairness may develop over the sequence of peer assessment activities. On a positive concluding note, some interviewees remarked that being involved in intra-peer assessment helped them to develop an understanding of the ethics involved in peer marking and that it "teaches students to avoid bias" (MR).

---

[4] The assumption, that opportunities for unfair marking will be minimised under those conditions, is, however, somewhat simplistic. As suggested in the earlier paragraph, the origin of some causes of unfair marking may be manifest before subsequent rehearsing and performing cycles indicated by the axial coding sub-category *Friendship marking is a problem*, described in Chapter Five.

6.8 The process model and band determined personal weaknesses

The research findings indicate that students can develop and improve their personal weakness attributes by revealing them to their band members and, importantly, agreeing that they may be used as assessment criteria. This is the key activity of stage three where, subject to students' agreement, their band determines their personal weakness attributes; it is an important activity that gives the process model its distinctiveness. Indeed, the quantitative findings suggest that bands, as an entity, can provide a useful contribution in this area. For example, quantitative data analysis at an early stage in the action research cycles indicated that band members consistently agreed with each other when awarding marks for their members' self-selected personal weakness attributes. Would they also agree with each other over the identification of their members' personal weaknesses?

To recap, the initial idea involving asking the bands, themselves, to decide personal weakness attribute criteria for each of their members, crystallised during a small number of interviews with students[5]. Students agreed that those who spoke with the greatest authority, in deciding which personal attributes might be improved in rehearsing for individuals, was their band.

As emphasised elsewhere in this chapter, the principle of agreement is integral to the model. It is important that the use of band determined personal weakness and the particular personal weaknesses that the band chooses should be made with the full permission and agreement of the individual. In this study, I considered it essential that written agreement was sought from individuals in advance of rehearsing. Bands, for example, might have chosen a personal weakness that an individual considered was unable to improve or develop. Individuals may, understandably, object to attributes that they consider are unfair. For example, perhaps a previous intra-peer assessment experience involving a similar attribute in which the student felt s/he was wrongly assessed, or a suspicion that hiding behind a band determined personal weakness is an interpersonal issue, unconnected with rehearsing. Objections to particular band-determined personal weaknesses require the tutor to be responsive and sensitive to the interested parties in seeking a resolution. There is little to be gained from insisting that students will be assessed on a personal weakness attribute

[5] See Appendix 7 (p.290).

without their agreement. Indeed severe damage might be done to the student-band and student-tutor relationships, as well as undermining the rationale for intra-peer assessment. Reaching agreements, revising personal weakness attributes and obtaining written permissions[6] are activities that are especially advisable for stage three of the model.

Difficulties will be resolved if they are seen to be done fairly. The social context is important: bands should discuss and formulate band determined personal weakness attributes together as an entity and present these, as a band, face-to-face with the individual. The overwhelming majority of students in my study accepted the personal weakness attributes that the band had provided with little hesitation. However, it was explained that if problems arose from the personal weakness attributes that had been determined for individuals, the band should revise or renegotiate those attributes contested by the student until agreement was reached. This safeguard is advocated for the model: queries about attributes can be quickly resolved via an amicable discussion with the band (perhaps leading to some alteration in the wording of the attribute).[7]

As remarked, in Chapter Five, the actual moment of receiving what their band adjudges to be their three personal weaknesses would, for many individuals, be a profound feedback experience.[8] Indeed, this kind of feedback was, for some, a maturing and revelatory experience; it can help you to see yourself as others see you. It enables a desire for the improvement of one's image, in the band's estimation.

Finally, the tutor needs to make a judgement regarding levels of trust among students and their bands, and also between students and tutor in stage three. If sufficient trust has developed, then *Group Attributes* may be discarded at stage four, in order to enable an increased focus on personal weaknesses, as previously explained.

---

[6] In conjunction with the *personal weaknesses* that the bands had generated in this study, I devised an agreement form (Appendix 9, p.305) to signify the student's acceptance of these and permission for them to be used as assessment criteria. Incorporating an agreement form supports transparency of the model.

[7] There were only two instances in this study when agreement could not be sought. In both cases, I considered it fair that the individuals concerned should have the option to self-select their *personal weakness/weaknesses* or, re-use the *personal weaknesses* that they self selected from the previous assessment. Each student agreed with the latter: repeating the personal weakness attributes as used previously

[8] Individuals' immediate acceptance and agreeing to the band determined *personal weaknesses*, as was typical in my research, also suggests that the student's self-awareness and band's perception of him or her were often 'in-tune' and confirmatory of each other.

Band determined personal weaknesses occupy a central position in the study, as explained in Section 6.4. They are a key part of the epistemological and conceptual underpinning of the process model that has developed out of the action research. Band-determined personal weaknesses more than any other aspect of the model, enable a focus and synthesis of the various strands of the enquiry. Figure 6.7 (p.222), illustrates this focus on band determined personal weaknesses involving the qualitative and quantitative data analysis, the rationale, the pre-requisites and the tutor's role. Specific examples are also provided for each of these areas of the enquiry. Double arrowheads linking related areas of the enquiry indicate a synthesis of this: for example, the tutor with pre-requisites and rationale; the tutor with rehearsing and performing cycles; quantitative data and qualitative data analyses.

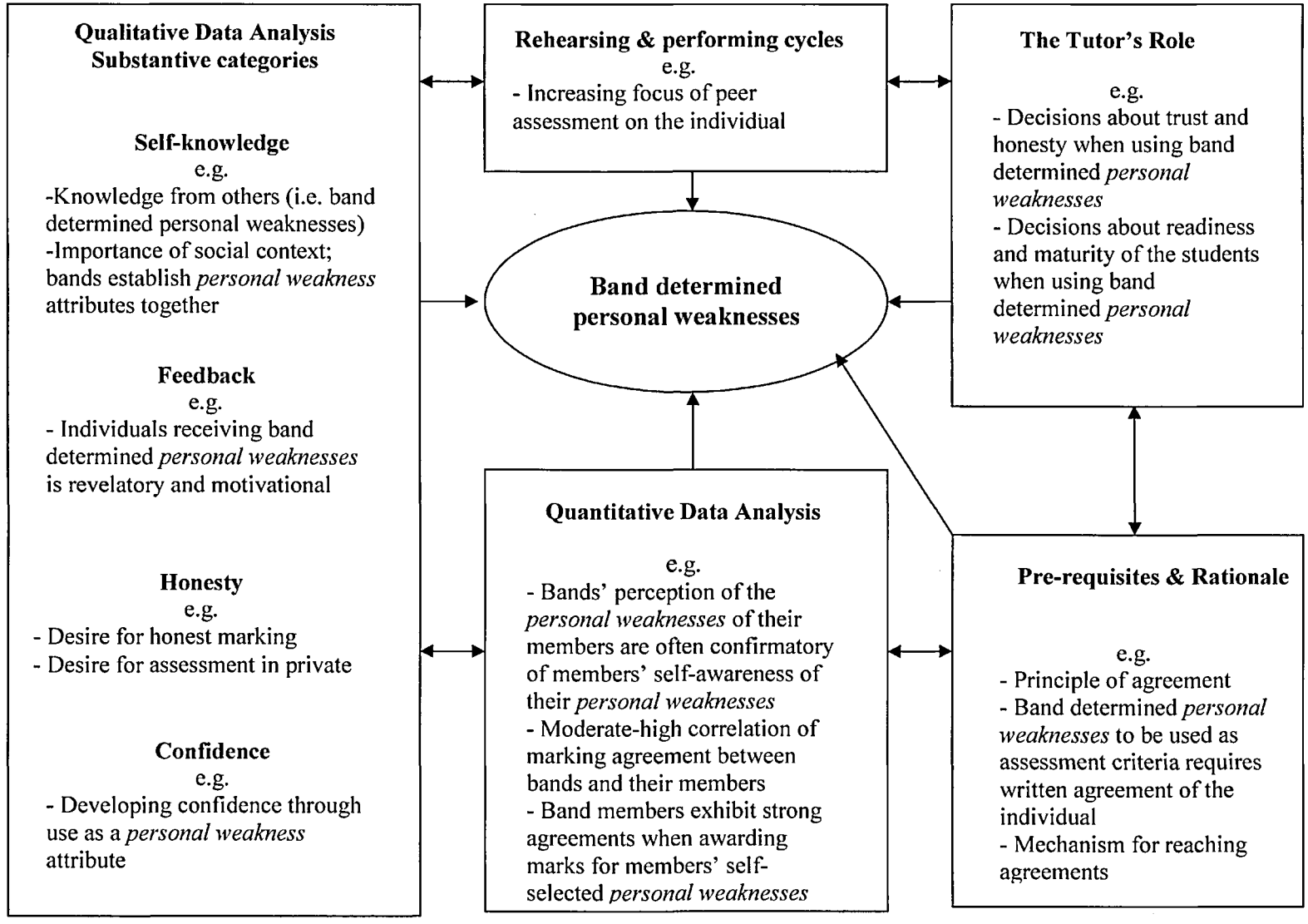


Figure 6.7 Focus on band determined personal weaknesses arising from the enquiry

6.9 Strengths and weaknesses of the intra-peer process model

This chapter has proposed a new process model for intra-peer assessment, founded on a graduated sequence of activities, which progresses student' learning over a realistic period of time, supported by a theoretical underpinning. At this point, it is appropriate that an examination of the strengths and weaknesses of this model is conducted. General problems such as unfair marking and weaknesses or strengths that may be considered generic to peer assessment will not be discussed here. Rather, this section identifies strengths and weaknesses that have specificity to the model.

In surveying its strengths, the model can be seen as founded on a number of pedagogical and paradigmatic principles. It offers, through its stages, a graduated process of engagement in intra-peer assessment that involves students developing awareness and knowledge of themselves and others. It is also flexible: there are a number of options available for progressing students on to the next step, depending on their readiness. Placing personal attributes at the centre of assessment criteria enables each student to develop and improve key attributes that are tailored to each individual and relevant for their own rehearsing and performing context. The model supports peer learning and the creation of knowledge about its participants, taking place within a constructivist-learning environment relating to band rehearsing and performance. It carefully prepares each student for activities involving band-determined *personal weaknesses*: accessing a distinctive knowledge about themselves, through allowing and agreeing to their band deciding their *personal weaknesses* for them. This activity provides an exceptional type of feedback, which enables individuals to develop their self-awareness through reflection, and to acquire knowledge of themselves and others in rehearsing and performing. Such knowledge seems to enable individuals to be more effective in rehearsal, leading to improvements to their band performances. In its use of personal attributes as assessment criteria, the model recognises the importance of careful preparation and support for learning. It is responsive to matters of trust, honesty, confidence, fairness and transparency, and embodies the principle of agreement.

In surveying the weaknesses of the model, a number of questions have already been raised For example, it was noted in Section 5.3.4 (p.203) that changes to a band's rehearsing strategy impacted adversely on a particular individual in their attempt to demonstrate

improvements. There are two other concerns, however, which pose a potentially greater threat to the viability of the model. These are: proceeding through the stages when agreement is not obtained with every individual; deciding to progress to the next stage when a minority of the class may not have exhibited signs of 'readiness'.

Firstly, and despite careful preparation and support, together with the tutor's judgement that sufficient trust and honesty has developed, individuals may nevertheless exercise their right to opt out from intra-peer assessment processes: the principle of student agreement is enshrined in the model. Consequently, if a student does not agree to a particular intra-peer assessment process and alternative arrangements prove fruitless, this can impact on their remaining band members and may lead to a problematic duality of assessment procedures. For example, the tutor may face a situation where four band members have agreed to band-determined personal weaknesses and the fifth member has not agreed. Although there is a mechanism, explained in Section 6.8, whereby individuals, on failing to reach agreement about their band-determined weaknesses, may self-select personal weaknesses as an alternative, what if agreement to disclose personal weaknesses to others is not obtained at stage one? Facing this situation, the tutor might consider, perhaps, the use of intra-peer assessment and tutor assessment simultaneously.

The second concern involves students' 'readiness' and 'progression' from one stage to the next. Deciding whether students are ready to progress is a matter for the tutor, based on his/her experiential knowledge of the class. It involves a judgement, for example, about whether sufficient trust and honesty has developed in the previous stage, as explained in Section 6.2. 'Readiness' in the context of the model is not, unfortunately, likely to be uniform among all members of a band, let alone in the class as a whole. Although the tutor will have recognised signs of 'readiness' (for example, maturity, trust, honesty, interpersonal qualities) among a majority of individuals in the class, there may be a minority who need more time for this to develop.

Does the tutor accept the risks involved in moving all the class on to the next stage, despite having some doubts about a few? Ought the tutor delay progression to give more time for 'readiness' to appear in this minority and, in so doing, perhaps holds back the majority and deny them the experiences from which they might have benefited?

The model offers no easy answers to these and the other problems that were identified. It is, however, wise to anticipate such concerns and develop appropriate strategies and contingencies. These are issues that require the tutor to decide what is appropriate within their rehearsing and performing contexts.

In summary, this chapter presents a new and distinctive process model for intra-peer assessment that has developed from this action research study. It advances our understanding of this mode of assessment and it offers new possibilities for supporting learning. The process model involves students in a graduated approach to learning using intra-peer assessment comprising a sequence of four stages. Students and bands formulate their individual and group-agreed attributes, which are used as assessment criteria for providing feedback on their rehearsing and performing activities. These have an increasing focus on the individual, culminating in stage four, where group-agreed attributes may be discarded in order for intra-peer assessment to be focussed entirely on individuals' *personal weaknesses*. Feedback arising from this develops students' awareness and knowledge of themselves and of others during rehearsing and performing, leading to improved band performances. The process requires careful preparation, and this is integral to the model. This includes: developing trust and honesty between individuals, their bands and tutors; deciding on the 'readiness' of students for progressing to the next stage. There are a number of options within the model that provide flexibility in terms of 'readiness'. *Band determined personal weaknesses* are a key element of the model. These involve individuals agreeing to allow their bands, with whom they have rehearsed, to decide their *personal* weakness assessment criteria. The resultant feedback is a key knowledge outcome, supporting students in learning about themselves and others. In short, such feedback can help you to see yourself as others see you, and to know yourself as others know you.

The chapter concludes with an examination of the strengths and weaknesses of the intra-peer process model. A number of illustrations are provided throughout the chapter in order to support, clarify, summarise and contextualise the model. For example, Fig 6.6 (page 214) suggests possible associations between types of personal attributes arising from the data analyses and the process model; Table 6.1 (page 216) is a summary of the relationship between trust, honesty and the process model; Figure 6.7 (page 222) illustrates the focus on band determined personal weaknesses arising from the enquiry.

The culmination of the action research is represented by the creation of a new process model of intra-peer assessment for group work rehearsing and performance, appropriate for undergraduate students of popular music. Chapter Seven considers this enquiry across a variety of contexts and examines the wider implications of its findings.

Chapter Seven: The research in context

7.1 Examining the research through different perspectives

7.2 Implications of the action research for tutors

7.3 The curriculum context

7.4 The methodological context

7.5 The personal context

7.6 Summary of the contribution to original knowledge

Chapter 7: The research in context

The two preceding chapters occupy critical positions in this enquiry: Chapter Five presented an analysis and interpretation of the qualitative interview data and quantitative peer marking data; Chapter Six proposed a new process model of intra-peer assessment that was informed by these findings. Chapter Seven moves the enquiry from a micro level of engagement, represented by a concern for the processes of the model in relation to the empirical findings of the research, to a macro level, which examines the enquiry in wider contexts. The purpose of this final chapter, therefore, is to present an exposition of theoretical and contextual matters that are raised by this enquiry. It concludes with a summary of the original contribution to knowledge that this research has produced. This chapter, accordingly, is organised into the following sections.

7.1 Examining the research through different perspectives
7.2 Implications of the action research for tutors
7.3 The curriculum context
7.4 The methodological context
7.5 The personal context
7.6 Summary of the contribution to original knowledge

7.1 Examining the research through different perspectives

An earlier discussion about social constructivism and the relevance of this perspective for the nature of the enquiry can be found in Section 2.2 of Chapter Two. Assumptions based on social constructive epistemology, while acknowledging some overlap with cognitive constructivism, were thought to provide a paradigmatic foundation for the developing research. These assumptions included, for example, the following: the centrality and social setting of the rehearsing, performing and peer assessing experiences for creating knowledge; the tutor and research roles being interactively linked to such experiences and that knowledge was created as the enquiry proceeded.

Similarly, there are epistemological assumptions, implicit in the development of process model of intra-peer assessment, which are also thought to involve social constructivism. These included, for example, the following: deciding on students' readiness to progress on the basis of the tutor's experiential knowledge of them; supporting learning through

the provision of social contexts for formulating personal attribute criteria.
However, it needs to be considered whether a social constructivist perspective, alone, is sufficient in accounting for the findings arising from this research and for providing the pedagogical underpinning of the process model. Does social constructivism, as applied to this research, explain the changes involving personal attributes and learning behaviours that interviewees reported, and which were discussed in Chapter Five? Is social constructivism of relevance in explaining how and why personal attributes improved or did not improve? Does social constructivism offer an explanation of how students assumed their roles of formulating and ascribing meanings to others' *personal weakness* attributes? Before re-visiting social constructivist theory, in view of the completed action research, these and other questions suggest the possible relevance of at least two other major perspectives for the thesis: social cognitive theory and symbolic interactionism. The social cognitive learning theories of Bandura (1977, 1986), for example, which combine ideas derived from behaviourist traditions and social thinking, can illuminate important constructs when applied to this study. Symbolic Interactionism, a major sociological perspective that studies, among other phenomena, how meaning is constructed through face-to-face interaction, similarly has significance for this enquiry. Although different to social constructivist theory, symbolic interactionism, nevertheless, shares many of its beliefs.

The purpose of these next two sections, therefore, is to illuminate the enquiry through the application of other major theoretical perspectives and, in the course of this, to search for alternative explanations. Following this discussion and analysis, social constructivism is once more revisited in Section 7.1.3.

7.1.1 Social Cognition

Bandura is an important and influential contributor in the field of social theories of learning. Although many might associate social learning theory with behavioural theories, Bandura (1986) describes his ideas as 'social cognitive theory'. Despite being beyond the scope of this chapter in presenting a detailed account of his theory and its application, it is suggested, nevertheless that many of its constructs do have a relevance to the enquiry. He maintains that people possess self-beliefs and that 'what people think, believe, and feel affects how they behave' (Bandura, 1986, p.25). In an overview of

Bandura's ideas, Pajares (2002) points out that the importance for also working together on shared beliefs about their capabilities and common aspirations; this has relevance for the formulation by a band of their group attributes[1]. Pajares', continuing the overview, discusses the importance of human capabilities such as self-reflection and motivation for social cognition, which are similarly relevant to this study.

At the centre of Bandura's social cognitive theory is the principle of self-efficacy belief. In this, it is asserted that unless people believe that their actions can lead to improvements, they have little incentive to act on it. In this study, self-selected personal attributes, such as personal weaknesses, may be illustrative of an individual's self-efficacy belief. For example, selecting *confidence* as a *personal weakness* attribute would demonstrate an individual's belief that s/he possesses the capability of developing and improving it. Conversely, that same individual might not self-select certain other personal weaknesses that they consider to have because of doubting their efficacy to mount a similar effort for those attributes.

Bandura identifies a number of sources from which individuals form their self-efficacy beliefs. These can be described as: mastery experience (meaning previous performance); modelling (meaning the observation of others); social persuasions (meaning the feedback and judgements of others); somatic and emotional states (such as anxiety and mood). Of these, mastery experience would have particular relevance for individuals in reflecting on their personal attributes in a previous rehearsing and performing experiences, and a few examples of this appear in the interview data. The occurrence of modelling arising in the action research, however, is less easy to demonstrate. Interviewees did not specifically cite instances of being influenced by, or of attempting to model their own attributes on those of their band members. There is a clear link, however, between peer assessment and 'social persuasions', both in marking feedback as well as for individuals receiving their band-determined *personal weaknesses*. With regard to somatic and emotional states, there is little concrete evidence arising from the study to explain that these had a bearing on the efficacy of individuals. Although 'bias' and other expressions of unfairness or tension arising from peer assessment, as evidenced in Chapter Five, were manifest, it is not apparent that these affected the self-efficacy belief of an individual.

[1] Group-agreed attributes, as explained previously.

Self-efficacy appraisal occurred during activities involving self-assessment. The under- and over-estimations exhibited by shooting stars and free-riders respectively between their self assessments and those of their band has implications for misjudgements about self-efficacy. Such errors of judgement in self-efficacy, Bandura argues, require to be checked periodically in order to ascertain the effect of further experiences on competence. It suggests that a wide variance between, for example, a self assessed and band assessed personal weakness attribute, requires a reselection of that attribute in a subsequent intra-peer assessment, in order to improve the self-efficacy judgement of the individual. This, in fact, actually occurred with regularity between cycles when, for example, an individual's self-selected personal weakness attributes that s/he formulated in one cycle also appeared in the next cycle as band-determined personal weaknesses.

Social cognitive theory provides, even in this basic account, an interesting perspective towards offering an explanation of how and why students' personal attributes improved. Through its principles of self-efficacy, it illuminates environmental factors and cognitive processes whereby individuals develop self-awareness and motivation for improving their personal weaknesses. Social cognitive theory also emphasises the importance of observation between individual band members, during their rehearsals, and learning or modelling from this. What Bandura's theory might find difficult to explain in this study, however, is the development of a moral knowledge including values such as honesty and trust. Neither is it clear how a social modelling environment and self-efficacy goes beyond task performance, skill acquisition and behavioural outcomes. Can these also account for knowledge creation, described in Chapter Five, which arose, for example, from the metacognitive competencies of band members during the complex social process of formulating the *personal weakness* attributes of others?

7.1.2 Symbolic interactionism

A symbolic interactionist would comprehend the research in very different terms to that of a social cognitive theorist. Of fundamental importance for symbolic interactionists are the meanings, in this study, that are given to personal attributes by band members and the social processes involved in these. Symbolic interactionist theory raises several profound questions about the research. How do individuals self-selected personal attributes and what meanings would the individual and the band members give to these?

What are the differences, symbolically, between group attributes and personal attributes?

In addressing this last question, group-agreed attributes were developed in a face-to-face band interaction. It would have symbolic meaning surely for the band as an entity, being based upon what its members thought was important for them. Moreover, formulating group attributes would involve meta-cognition that recalled and synthesized previous accumulated experiences. Deciding on a descriptor (for example, *commitment*) to best represent group-agreed attributes would require a negotiation activity also.

Self-selected personal attribute criteria may, in a symbolic interactionist approach, be considered as a means whereby the *self* or *identity* is presented and constructed. Stage two of the intra-peer assessment process model presents the first occasion for students to engage in this activity of constructing meanings about themselves that are symbolised in personal attributes. This process might involve an awareness of one's own image. Deciding on a descriptor to represent this (for example, *versatility*) is a process that, for example, may derive from class discussion, or arise from the complexity of sociolinguistic meanings that are comprehended by the individual band member. It would, I believe, be a highly cognitive activity. Indeed, of particular interest in this enquiry are the thinking processes that occur between the following two points: being presented with the activity of formulating a personal attribute and through self-awareness and cognition crystallising it symbolically; deciding a descriptor that best represents the symbolic meaning of the personal attribute.

Symbolic interactionists often use metaphors borrowed from dramaturgy in order to explain situations, including, for example, 'actors', 'role-making' and 'role-taking'. A simple analogy can be made using these metaphors with this enquiry: the actors and actresses are the students and myself; assessed students are participants in *role making;* those determining the personal weakness criteria for others are engaged in *role taking* as they are taking another person's perspective. Chapters Five and Six each identify honesty and trust as important themes arising from the research and, in the development of the intra-peer process model. In the same way, the actors in the rehearsing and performing cycle assume roles requiring trust (for example, in allowing peers to identify one's personal weaknesses) and of honesty (for example, in their marking). Similarly,

group-agreed attributes can be understood as symbolising negotiated constructs that signify a bond of trust. As discussed in the previous chapter, group-agreed attributes have importance in the intra-peer assessment process model for developing trust among the members of the band.

Symbolic interaction theory regards social constructed relations as temporary and dynamic. In the same way, band determined personal weaknesses, for example, might be considered as socially constructed symbols representing a particular moment of interpersonal interaction in the rehearsing and performing life-world of a band. During rehearsal interaction, new relationships and social constructs of a temporal and transient nature evolve which may cause strain upon the initial personal attribute criteria that was formulated. In other words, that which was originally created and symbolised as a part of an individual's self (for example, a personal weakness attribute) would become now separated from him or her because of the changing relationships and situations of the band and of individuals belonging to it. Such a separation would also suggest personal weakness attributes existing independently from those who constructed them, a scenario which symbolic interactionists would find difficult to accept. If so, then this raises the question of whether personal attribute descriptors should, out of necessity also evolve, because of the readjusting social processes in the rehearsing life-world of the band. It was an issue that arose from the interview data in Section 6.3.4 and one that attracted brief speculation in Section 2.3.5. Attempting to resolve this difficulty, however, might lead to tensions occurring across three dimensions: the real-world practicalities of changing the criteria midway through an assignment, injustice to individuals caused by not changing, and institutional regulations governing assessment procedures.

In conclusion, a symbolic interactionist approach illuminates many aspects of the enquiry with an easiness and naturalness. It offers a sociological insight at a band rehearsing level and adds considerably to an understanding of the meanings attached to students' personal attribute criteria as they evolve during their rehearsing interactions. Although it is less able to focus on macro-structures such as the intrapeer assessment process model, symbolic interactionism enriches the contextual and theoretical understanding of the research at a micro-level. Figure 7.1 (p.234) summarises a symbolic interactionist perspective of the enquiry.

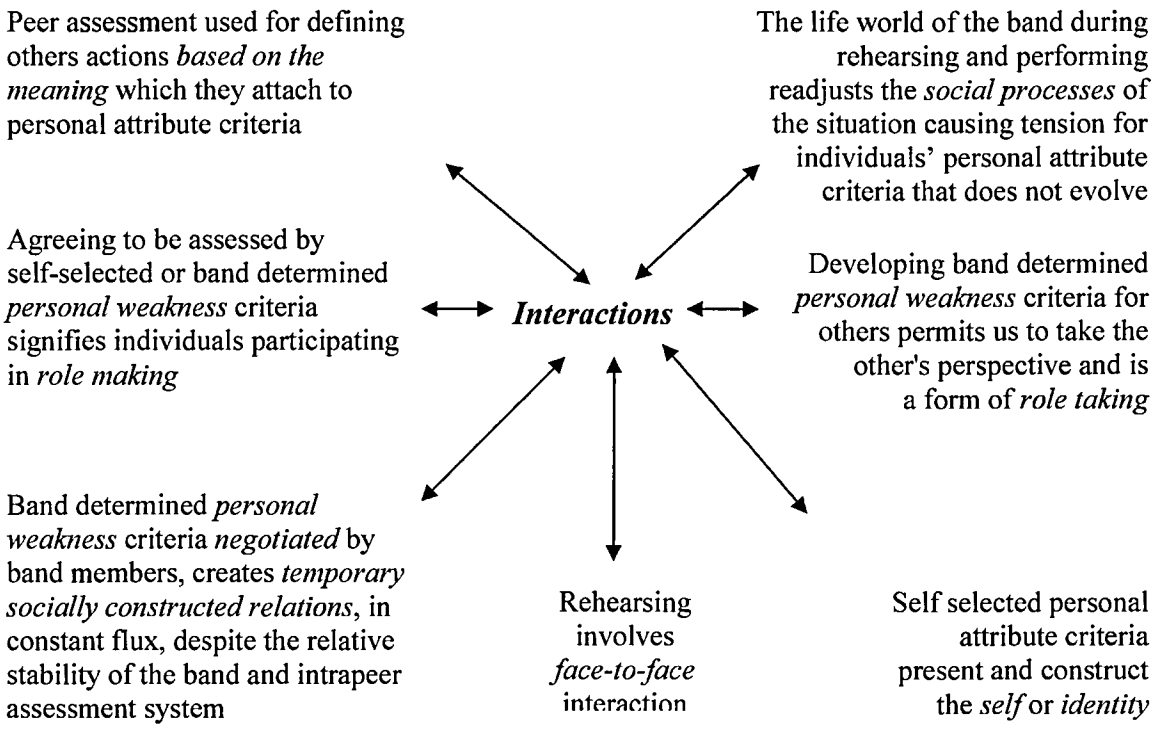


Figure 7.1 A symbolic interactionist perspective of the enquiry

7.1.3 Revisiting social constructivism

At this point, it is appropriate to revisit social constructivism and its place in this enquiry. As explained in Chapter Two, *Methodological Considerations*, the investigation adopted a social constructivist perspective and this section offers a critical examination of social constructivist theory in the context of the completed action research.

Social constructivism acknowledges the importance of the tutor in facilitating learning and knowledge creation (Vigotsky, 1962; Bruner, 1986; Fosnot, 1996). The various decisions that, as the tutor, I made in this enquiry, particularly in the process aspects (e.g. timing of activities, methods of intra-peer marking, etc) could have influenced the kinds of knowledge that might be created. In the training activities, as well as in rehearsing and performing cycles, tutor-students interaction involved thinking about a number of assessment-related activities - particularly focussing upon the activity of

developing personal attributes assessment criteria. There is, in some sense therefore, a duality of knowledge construction between the tutor (myself) and students in such an interaction. Nystrand (1997) identifies a number of specific methods, which may help to promote a duality of knowledge construction. Among these are 'peer response conferences' in which students meet in small groups to review each other's work; this parallels the activities in which bands devise their group attributes and band determined personal weakness criteria.

Similarly, tutor and students, in their developing relationship, face a number of decisions during rehearsing and performing. Bands are presented with decisions or choices arising from the intra-peer assessment process model. For example, they decide upon their songs, sometimes the group membership, and whether to intra-peer assess privately or collaboratively. Of particular importance, however, are the decisions that they make about personal attribute criteria and marking feedback.

Creating and making choices are important activities for tutor and students in their interaction with each other. This relationship can be likened to the model of Wells (1999), as discussed by Skidmore and Gallagher. Wells advocates a democratic setting in which groups of students have considerable latitude to choose specific topics and methods of enquiry. The relationship between tutor and students, however, is 'not a dialogue between equals' (p. 242). As a tutor, I had overall leadership responsibility for deciding or facilitating the various options and choices that were required during the action research. However, this relationship was also dynamic, and changing at various points in the rehearsing and performing cycle; this aspect is explained in the next section, where it is linked with trust and honesty. In bringing together constructivist views of knowledge, rehearsing and performing interactions, and the relationship between tutor and students, it is possible to develop an epistemology arising from the enquiry. Figure 7.2 (p.236) summarises the relationships that are involved in interactions during rehearsals.

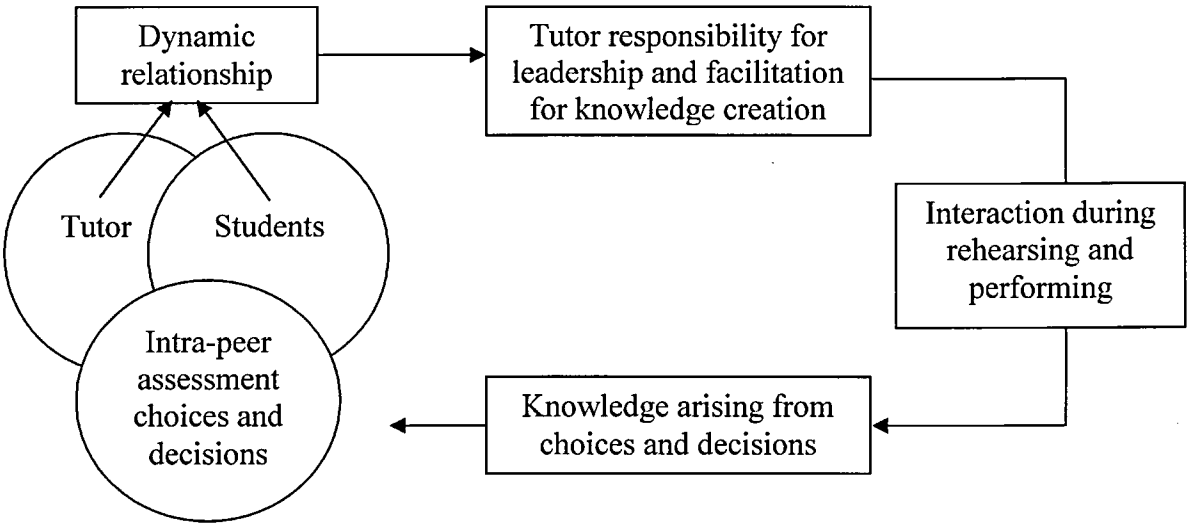


Figure 7.2 Epistemology of rehearsing interaction, choices and decisions

The seeking of intersubjective understandings through interaction typifies constructivism. In seeking these, and recalling aspects of symbolic interactionism also, rehearsal interaction may foster the development of students' understanding of a shared language of personal attribute criteria in a rehearsing context, which, in part, is also non-discursive. In this sense, it is proposed that non-discursive knowledge is created and shared by students arising from their understanding of the meanings of personal attributes within rehearsing contexts.

Providing opportunities for interaction, leading to knowledge creation, has implications for the time that is available for this. In this enquiry, the majority of tutor-student intrapeer assessment interaction time occurred within the rehearsing and performing cycles. Tutor-student interaction additionally took place through the interviews. The greatest amount of time available for student-student intra-peer assessment interaction arose within the rehearsing and performing cycles. Student-student interaction also occurred during the formulation of group attributes, band determined personal weaknesses and collaborative intra-peer assessment settings. Figure 7.3 illustrates interactional opportunities for knowledge creation.

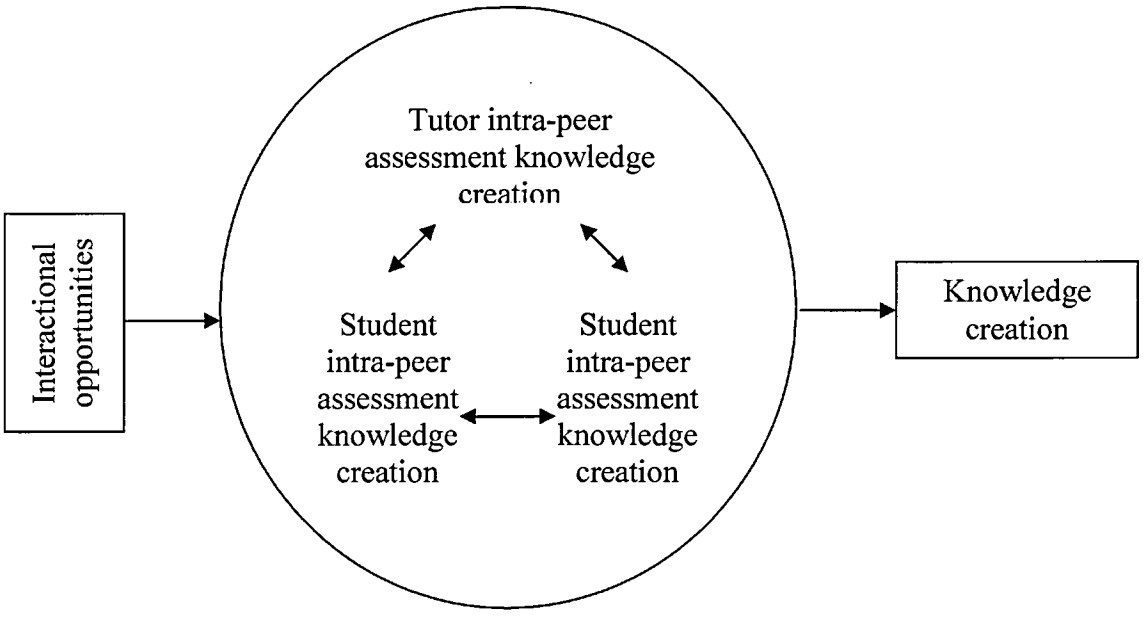


Figure 7.3 Interactional opportunities for knowledge creation

At this point in the chapter, certain moral and pedagogical considerations enter the discussion. Constructivism involves a shift in understanding away from knowledge as objective to an understanding of knowledge as value-laden and teaching as a moral activity (Gallagher, 2004). Adopting a constructivist perspective towards HE popular music students and their band rehearsing required an epistemological shift in assumptions about the nature of knowledge. This impacted on my dual roles as tutor and researcher in my interaction with the considerable numbers of learners that I encountered during the course of this action research.

Chapters Five and Six explain the rehearsing and peer assessment experiences from a moral perspective involving, for example, honesty, trust and respect. A moral dimension emerged from the life-world of the bands, their intra-peer assessments, and intra-student and tutor-student interactions. Students, their bands and the tutor learned together in an experience that developed their 'dynamic relationships' in an intrapeer assessment context. The moral context for learning, together with a 'dynamic relationship' may have resonance with constructivism; indeed, this 'dynamic relationship' can be linked with moral values such as honesty and trust. Consequently, the discussion in this section concludes with an examination of such a proposition.

Earlier in the chapter, it was suggested how a symbolic interactionist might explain the meanings that students attributed to their actions. It is to the meanings of honesty and

trust that the concept of 'dynamic relationship' may relate. The concept is used in the sense that such meanings may change for the 'actors' involved in the intrapeer assessment process model. The relative emphasis on and meanings attributed to either honesty or trust can be dynamic and in flux at or during each of the stages in the model. An analogy may be made here to Young (1991), in his paper on symbolic interactional theory, where he regards honesty and trust, essential to all symbolic activity, as non-linear psychological processes that produce non-linear effects. For example, the level of honesty and trust that students have at the starting points of the model is likely to be the result of previous interactions with others; this will vary from individual to individual. Levels of trust will also change within a band if, for example, the presence of *free riders* requires a higher level of trust in order for cooperation to occur.

Figure 7.4 (p.239) illustrates the complexity of these changes in meaning and emphasis of honesty and trust in the dynamic relationship between students and tutor arising from the model. At stage one, for example, there is an emphasis on trust whereby students are trusting of the tutor's understanding of them and of the system. Correspondingly, tutors place trust in the students and faith in the system. This changes during stage two where, for example, honesty and trust have meanings that involve personal attributes and for the marking of these. For stage three there is an emphasis on trusting bands to determine personal weaknesses appropriately and sensitively. Trust extends in two directions: bands, empowered with this responsibly, trust themselves to formulate band determined personal weaknesses with appropriate care; individuals trusting bands to determine their personal weaknesses with integrity. At stages two and three, honesty and trust have meanings related to the kind of feedback that is given or received. It would be unwise to assume that honesty and trust always have the same meanings for the 'actors'. Differences of meaning can arise, perhaps, from inequalities in the relationship between students and tutors involving, for example, the tutor being regarded as a figure of authority.

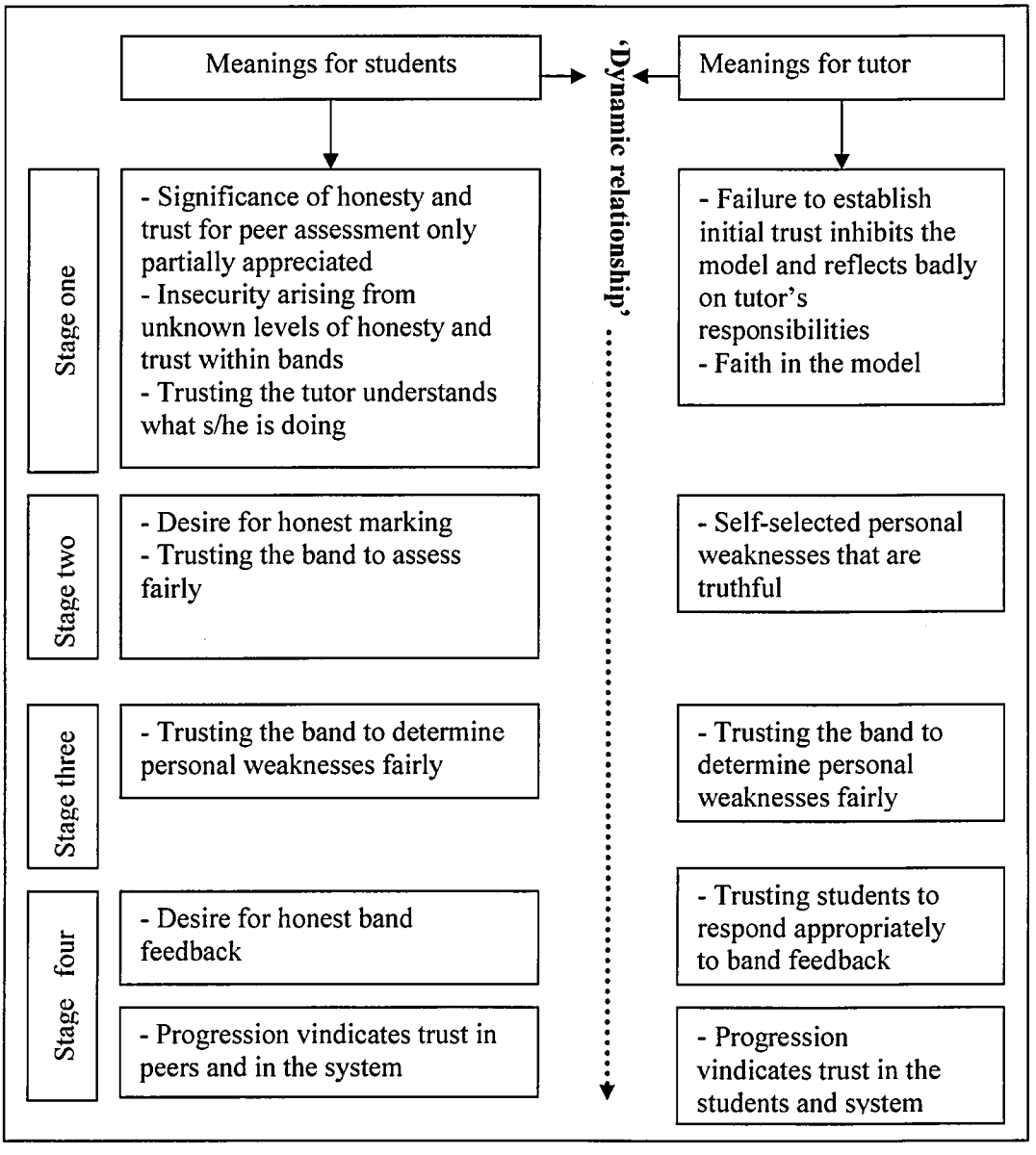


Figure 7.4 Meanings of honesty and trust in the 'dynamic relationship' involving students, tutor and the intra-peer assessment process model

To summarise, three theoretical perspectives of the enquiry have been offered towards providing explanatory foundations to the enquiry: social cognitive theory, symbolic interactionism and social constructivism. Each tradition illuminates important dimensions: for social cognitive theory it is self-efficacy belief that develops individuals' capability for developing their learning; for symbolic interactionism it is the meaning that is attributed the 'actors' personal attributes in intra-peer assessment that explains their actions; whereas for social constructivism, it is the peer-peer and peer-tutor interaction that creates a particular kind of knowledge for the participants, which arises from the choices and decisions of intra-peer assessment, a part of which is moral knowledge. Figure 7.5 suggests, at a simplified level, a possible synthesis of these perspectives to the enquiry.

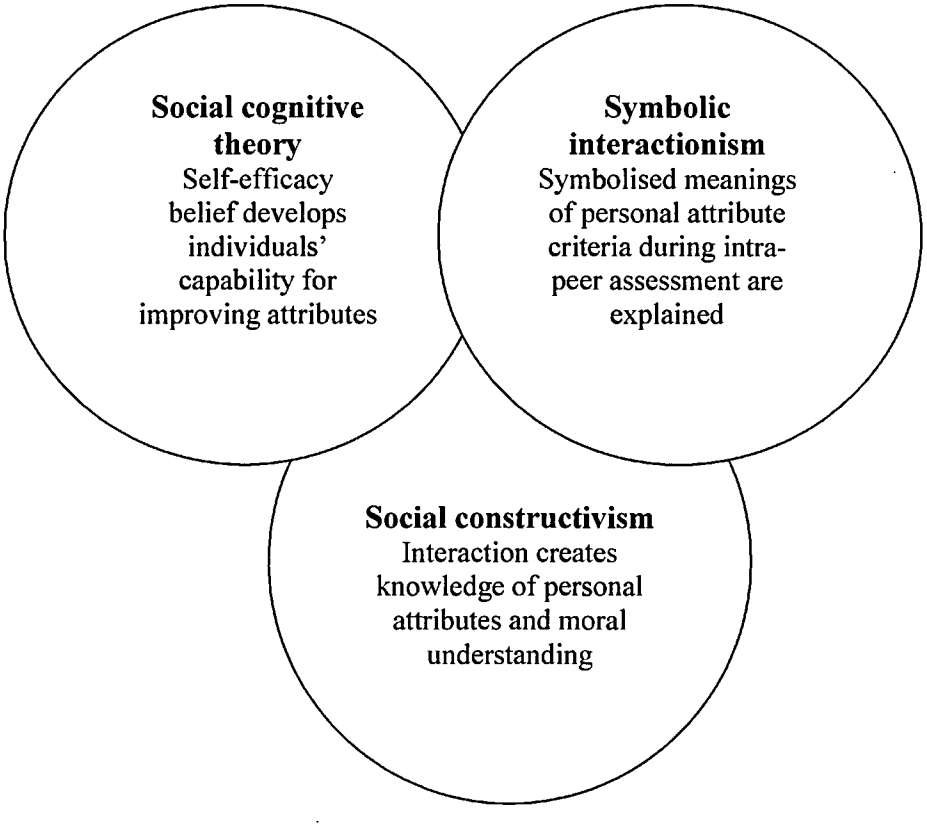


Figure 7.5 Possible syntheses of theoretical perspectives for the enquiry

Postulating an interconnectivity of these three traditions for the enquiry suggests the following: social cognitive theory may provide an explanation for students' behaviour in response to peer assessment, symbolic interactionism may explain meanings that are symbolised by students in the formulation of personal attributes, and social constructivism can provide an interactional context for the creation of knowledge. Does

such interconnectivity adequately explain how students meaningfully engaged with intra-peer assessment in the rehearsing and performing cycles, as they certainly appeared to? Is the provision of interaction contexts for rehearsing and performing, by itself, sufficient to allow knowledge to be created? Although offering an account of how self-efficacy belief develops individuals' capability for developing their learning, does social cognitive theory adequately explain how students come to comprehend and engage with it in the first place? Even though symbolic interactionism offers a compelling explanation of the meanings that are symbolised by personal attributes, does it adequately explain the contextual activity that provokes such meanings to become manifest and communicable? Perhaps there an intrinsic passivity about each of these viewpoints, which does not adequately account for students' initial comprehension of the complex worlds that are connected through this enquiry: rehearsing and performing, intra-peer assessment, group work, personal attributes, popular music. There may be a phenomenon occurring during intra-peer assessment that transcends, for example, the interactional band rehearsing context. The previous discussion about honesty and trust involved the concept of the 'life world of the rehearsing band'. It is to this and other 'worlds' that the present discussion now turns, in the search for an explanation of how students begin to comprehend and engage with intra-peer assessment.

7.1.4 Worlds and languages of the enquiry

Ideas derived from philosophy of language and meaning theory may provide a useful contribution at this point. Humans are able to make sense of the world and to comprehend their existence through, for example, the development of language in its various manifestations. As well as the language of science, of number, and of observation, there are countless non-discursive languages such as that of facial expression, gesture and physical appearance. A musical example of the language of gesture may be observed in a conductor, in which the gestures expressed through his/her hand movements communicate meanings for the choir, orchestra or band. Music is defined by Aspin (1984) as belonging to a language of organised sound. Following Wittgenstein (1953), such languages and worlds only become meaningful and communicative to those who have learned them and who can comprehend, for example, their syntax and inner logic. This proposition also recalls the earlier discussion in Chapter One (p.37) about the idea of there being a 'community of knowers' (Lave and Wenger, 1991; Rømer, 2002).

It is in relation to the particular kinds of worlds intrinsic to the subject matter of this thesis (popular music, groupwork and peer assessment) that, to enable band members to learn such languages, I am proposing the idea of 'active engagement' in order to explain entry into those worlds. This enquiry involves a complexity of these, including, for example: the interpersonal world of the rehearsing and performing band; the cognitive and metacognitive worlds of personal attribute awareness; the evaluative world of intra-peer assessment; the moral worlds of honesty and trust; and the aesthetic and cultural worlds of popular music. I use the term 'active engagement'; a construct that can be compared with Peters' (1965) notion of education as a form of initiation, to express a state of involvement. Simply being passive in the interactional setting of a rehearsal is not enough; I am suggesting that in order for students to comprehend the complexities of such worlds that are suggested in this enquiry, 'active engagement' is also required.

Occasions where 'active engagement' with these worlds in the enquiry may be identified include; face-to-face negotiation of group attributes, formulating self selected attributes; establishing band determined personal weakness attributes[2]; engaging with intra-peer marking; deciding about allowing band members to identify one's own personal weaknesses; making judgements about the level of honesty among band members. The central focus in the facilitation of 'active engagement' is, of course, the tutor. One example of successful active engagement in the world of personal attribute meaning can be illustrated by my not recalling any instances of an individual enquiring, 'what does [e.g. *commitment*] really mean?' Contrarily, free riders' over-estimation of their personal attributes that was found in the data analysis may be explained by their inability to engage with the language and meanings in the evaluative 'world' of intra-peer assessment.

Thinking about instances of 'active engagement' into these worlds may help towards explaining how students made sense of their involvement in intra-peer assessment. In developing an epistemology for intra-peer assessment and pedagogy for the process model, the 'active engagement' paradigm might provide the connection with, or the conceptual 'glue' between, worlds and the learning that may arise from them. Although it is not possible to provide a more detailed exposition within the confines of this

[2] The frequency with which individuals and bands selected *personal weaknesses* attributes related to, for example, a lack of *verbal input*, *communication*, or *contributing ideas*, may also be interpreted as a sign that 'active engagement' was important for students during their rehearsing.

chapter, Table 7.1 offers an epistemology involving an 'active engagement' paradigm that can be thought of as being implicit as a consequence of this enquiry.

Table 7.1 Suggested epistemology of the enquiry relating to the 'active engagement' paradigm

Active engagement may enable individuals to acquire the 'language' of the following 'worlds'					
Examples of the 'worlds' involved in the enquiry	*Active engagement* into aesthetic and cultural world of popular music	*Active engagement* into interpersonal world of band rehearsing	*Active engagement* into metacognitive world personal attributes criteria	*Active engagement* into moral world honesty and trust	*Active engagement* into evaluative world of intrapeer assessment
Examples of *active engagement* in the enquiry	Listening, analysing and discussing subject matter of popular music	Communicating with other band members in interactional setting of rehearsing and performing	Negotiating, face to face in developing group attributes; identifying self-selected *personal weaknesses*	Discussing band determined *weaknesses* and implications for agreeing to the band deciding these	Judging personal attribute criteria through giving and receiving marking feedback
Examples of claims for *active engagement*	Learning discursive and non-discursive or 'tacit' languages and knowledge Understanding of symbolic meanings of personal attribute assessment Developing self-awareness and creating self-knowledge Creating knowledge about others Creating and sharing *a posteriori* knowledge				

7.2 Implications of the action research for tutors

The role of the tutor was central in determining the kinds of student experiences that were possible in this investigation. This section, therefore, discusses the action research and examines the implications of this for others who may be considering introducing peer assessment along similar lines to those described in this thesis. In Chapter Two and elsewhere, I referred to the necessity for my development of what was described as 'sufficient confidence and understanding' of peer assessment. Having now completed the study, it is possible to suggest what sufficient confidence and understanding might comprise; this is indicated in the following points.

- Clarity of the peer assessment rationale; for example, using peer assessment to support learning.

- Clarity of the context of the activity, as this defines the kinds of knowledge and learning that may be created during peer assessment. Group rehearsing, for example,

involves an interaction context involving the interpersonal life world of a band. The creation of personal attribute knowledge through intra-peer assessment can become an objective in this activity.

- Clarity in and demonstration of transparency; for example, seeking agreements with students also encourages their *active engagement* in key activities that involve establishing and agreeing to their personal attribute assessment criteria.

- Clarity of the rationale for decisions that are made by the tutor; for example: conducting intra-peer assessment in private; introducing band-determined personal weakness criteria.

- Clarity in the preparation and support of students' engagement in peer assessment; for example: training activities; revisiting the rationale and discussing this during the stages of the process model when appropriate.

- Clarity about tutor moderation; for example, explaining the availability of this and safeguards for students together with a discussion about students' moral responsibilities.

Of particular value, towards the acquisition of confidence and understanding, is an ability to be flexible and responsive to the diversity of the popular music student body and their bands from year to year, cohort to cohort. I believe I developed these qualities through actively engaging with and sharing the student bands' life worlds during their rehearsing. From these shared experiences, I constructed a particular kind of knowledge about the students and their bands. Understanding and making sense of this knowledge allowed me to make informed decisions about them concerning the intra-peer assessment process model. The following points illustrated the range of decisions that I was required to make.

- The appropriate tutor/researcher balance; for example: interaction in the rehearsing life world of the bands; monitoring the intra-peer assessment system and the impact on tutor time and other resources.

- Deciding on the readiness of students to progress to the next stage in the intrapeer assessment model.

- Where and when to seek students' agreement of the intra-peer process and the agreed level of transparency; for example, whether to identify each student's assessments of their peers when conducting assessment in a private setting.

- How to develop or maintain levels of trust and honesty.

- When to tackle difficulties and disputes involving student-student, student-band and student-tutor relationships.

- The degree of risk-taking for each rehearsing and performance cycle appropriate to the particular student cohort.

- How to address the diversity of band rehearsing and performing experiences that were typically encountered across popular music cohorts. For example: differences in students' maturity, maintaining the appropriate tutor-student relationship, maintaining the effective functioning of bands; addressing problems of dysfunctional band members.

Tackling each of these was a steep learning experience, requiring me to 'let go' of my ownership of assessment and other areas typifying tutor power, in which I was previously accustomed. The shift away from the traditional power relationships between teachers and students (Skidmore & Gallagher, 2005) was not, in retrospect, especially challenging for me, as this change resonated with the deeply held values that I described in Chapter Two. Somewhat more difficult, however, was the necessity to develop a combined tutor and researcher role that would allow my experiential engagement with the student bands while simultaneously maintaining appropriate tutor detachment.

7.3 The curriculum context

The outcomes of the research enquiry have a number of implications for undergraduate popular music curricula. Peer assessment, as a technique for assessing students, can often be seen simplistically; often, it is employed as a means to an end: the marking and grading of work. In this thesis, one can hardly fail to notice the complexities that were

encountered, involving its many-sided approaches (the intra-peer enquiry, the inter-peer enquiry, the process features enquiry) explored during the action research.

Such complexities, however, ought not to deter curriculum designers from incorporating this mode of assessment, where appropriate, into their schemes. It may be that this particular mode of assessment, offers the most appropriate kinds of experiential learning that the course sets out to provide.

For example, in devising learning outcomes for their courses, most music educationists would certainly consider the development of students' personal attributes as important within their teaching and learning philosophy. Indeed, as well as reflecting the increasing diversification of assessment across the Higher Education sector, there is also a growing acceptance of student-centred methods of learning, and of courses that address the needs of its learners through an apparatus that is uniquely tailored to the individual.

Turning to the topic of course design and the implications for undergraduate group performance assessment, of particular interest are the changes involving the kinds of personal attributes that were used as assessment criteria. For example, the final cycles of the research relied on the use of individuals' personal weaknesses criteria and little else. These later cycles reflected my increasing interest in developing and improving students' personal attributes through a focus on the individual rather than as a member of a group, and their band determined personal weaknesses. This aspect of the enquiry can be portrayed in terms of changes in balance: between group and individual attributes and, between personal weaknesses and other intra-peer assessment criteria, such as personal strengths or tutor imposed criteria. Course designers, who regard personal attributes as a central concern within their course learning outcomes and who also recognise that peer assessment may be a useful assessment technique in this area, should strive to achieve a balance of criteria. For example, developing mutually agreeable rehearsing objectives for a band and maintaining its cohesion suggests a preponderance of group attributes. An emphasis on developing cooperation and teamwork within group rehearsals, rather than a summative performance, suggests an assessment structure based purely on personal attribute criteria appropriate for the rehearsing. Appropriate criteria may include personal attributes involving, for example, *negotiation*, *conflict management* and *communication*. Such qualities, however, as noted

by Bryan (2006) and discussed in Chapter One, although often present, are rarely a focus for assessment because of the emphasis usually placed on product or performance. Several types of personal attributes were explored in this research in order to focus on particular activities in order to achieve certain objectives or goals. Table 7.2 summarises these types of personal attributes and the implications of these for supporting learning and setting goals, which may suggest possible applications for group work and intra-peer assessment across other curriculum areas.

Table 7.2 Types of personal attributes, focus of learning and goal

Type of personal attribute	Focus of learning	Goal
General: intra-peer assessment system		
Employing *group attributes* as intra-peer assessment criteria	Agreeing on common values and objectives shared by the group	Improved group bonding and cohesion; building group trust
Employing *individual attributes* as intra-peer assessment criteria	Developing awareness of individual needs in group interaction	Improved interpersonal group working attributes
Specific: intra-peer assessment criteria		
Employing an holistic *contribution to rehearsing* criterion	Contribution to the group	Improved group bonding and band cohesion
Employing self-selected *personal weaknesses* as intra-peer assessment criteria	Developing self awareness and reflection	Individuals identifying and agreeing to these being used as criteria; feedback leads to self-knowledge and improvement; honesty
Employing band determined *personal weaknesses* as intra-peer assessment criteria	Developing awareness of others, negotiation and formulating feedback	Group trust; feedback leading to self-knowledge, improvements and honesty

This thesis may have implications, in particular, for areas of the curriculum that are characterised by creativity or group work. Similarly, it is relevant for programmes in which the assessment strategy includes an element of intra-peer assessment, or for course learning outcomes that involve the development of personal attributes. Most disciplines within Higher Education incorporate significant amounts of group work. Indeed, colleagues from Performing Arts, Dance, Drama, Communication, Media, Business, Management and Education, in particular, might recognise opportunities for intra-peer assessment involving personal attribute criteria relevant to the learning contexts that are appropriate to each of these disciplines.

Future developments might involve exploring the use of intra-peer assessment involving personal attribute criteria across interdisciplinary contexts. Such investigations might also raise interesting questions about the transferability of personal attribute criteria from one group work learning context to another and, as explained in Chapter Six, this can be problematic.

7.4 The methodological context

The methodology of this enquiry, described in Chapters Two and Three, is already extensive and the purpose of this section is that of reflection: final thoughts and pulling together some of the many threads.

The overall aims of the research suggested an enquiry that would be both exploratory and focused. Exploratory, because it involved a number of investigations across several areas (e.g., peer assessment, group work, personal attributes, process features); focused, in the sense that it was situated in the context of the students over a five year period. It was important, in view of the extensive scope of this research indicated above, that considerable thought was given to its methodology. Indeed two chapters: *Methodological Considerations* and *Methods of Enquiry* were necessary in order to explain the epistemological and other bases of the thesis, and to describe the action research methods that were employed.

During the first two years of the action research, I did not fully comprehend the implications of the large amount of data, both qualitative and quantitative, which an investigation such as this might generate. The data collection was a substantial undertaking in terms of time and other resources. Analysing the extensive and very detailed quantitative assessment data had a similar impact. The interviews accumulated rich, 'thick' qualitative descriptive data and the analysis arising from it overlapped with the gathering of this information. The occurrence of overlapping data (i.e. fresh data appearing, qualitative or quantitative, before previous data has been analysed) is often characteristic of action research investigations as explained in Section 2.5.[3]

[3] I also experienced a similar phenomenon in the writing of this thesis in that its various chapters were written, refined and rewritten in a non-sequential and overlapping manner. Although at the time, I took the view that this was somewhat untidy, I now believe that it was entirely appropriate, perhaps inevitable, in an enquiry of this sort.

The adaptation of grounded theory to the analysis of the quantitative assessment data, explained in Chapter Three and Chapter Five, appeared fit for purpose. Indeed its procedures were tested very thoroughly due to the extensive amounts of data presented for analysis. I regard the work involving this adaptation of grounded theory as being practical and useful; it offers an original contribution to educational methodology.

If I were to repeat this action research study, I would consider the use of a diary with which to record my interactions with the *Performance Management* students. Having this record would, I believe, add a very useful observational dimension, especially for studying a chronological development of learning experiences over a sequence of band rehearsals.

7.5 The personal context

Trust, honesty, awareness, thinking, knowledge, confidence and interaction are among the key themes for all who participated in this enquiry. My dual role of tutor and researcher required me to think carefully about my interactions with the students. This included, for example, thinking about: how peer assessment, involving personal attribute criteria, can help to support learning; how to engage students in the experience; how to stimulate students' thinking about their personal attributes; how to support individuals at critical moments during the intra-peer process model; and how to decide on the readiness of students to proceed to the next stage.

I consider that this research, with which I have been involved since 2000, has had a profound personal impact. Engaging with this type of research activity has contributed to reshaping my thinking and my values as an education professional. I have been privileged to learn so much about the students and how they responded to peer assessment. I am grateful to the large numbers of students who have made this research possible.

This section therefore, offers a reflective and personal account of the research. I thought it appropriate to present this chronologically year-by-year, following Figure 2.1 (p.88) in Chapter Two and Table 3.2 (p.104) in Chapter Three. Figure 2.1 identifies the foci of my reflection during the action research (for example, *development of researcher's understanding of peer assessment; rehearsing and performance cycles: development of research questions; development of methodological framework*).

It is hoped that the following reflections, *post inquærere,* are helpful for others who may, themselves, be exploring similar approaches to intra-peer assessment. I believe that aspects of my own practice improved with each rehearsing and performing cycle, although this is for others to judge. If so, my improving practice may have had a positive impact on how students, in latter part of the enquiry, responded to the peer assessment.

Rather than providing a lengthy exposition for this section and a discussion of themes that may have already been partly addressed, I have opted to indicate key points using a bullet point format.

2000-2001

This was my first year of involvement with the action research. During these initial cycles, I began to develop an understanding of, and fascination with, the operational mechanics of peer assessment. I felt that the following points were important during this period:

- Developing my confidence (for example, a growing confidence in inter-peer panels to mark band performances accurately);
- Understanding about the use of peer assessment (for example, discovering the limitations of using simple *contribution to rehearsing* criteria);
- Fascination for process features (for example, conducting intra-peer assessment in private and in collaborative settings: conducted it either before or after students were aware of their band performance marks;
- Recognising opportunities for supporting learning (for example, assisting students in formulating their own performance criteria);
- Developing honesty and trust (for example, beginning to trust students as peer panel assessors);

2001-2002

The processes involved in using personal attributes as intra-peer assessment criteria. I found absorbing. The following were important for me during this period:

- Developing my experience of intra-peer assessment (for example, learning about facilitating the generation of group agreed and individual personal attributes).

- Recognising difficulties (for example, using category weighted criteria for inter-peer assessment because weightings specified in advance were not always appropriate or reflected in the performances).
- Belief in processes (for example, intra-peer assessment was best conducted before students were aware of their band performance marks).
- Responsibilities (for example, ethical concerns about effects of peer marking on band cohesion, free riders and dysfunctional bands).

2002-2003

I felt quite confident about using personal attributes as assessment criteria and began to think further about how they might be utilised in order to support individuals' learning. I became aware too, of the increasing importance for building trust among students, especially in relation to the peer panels. Important for me, during this period, were the following:

- Excitement in exploring novel ideas (for example, the process of generating personal weakness criteria; using formative and summative assessment of personal attributes within a single rehearsing and performing cycle).
- Learning about the importance of honesty and trust (for example, deciding whether sufficient trust existed in order to allow students to form their own peer panels).

2003-2004

It was at this stage that I decided to introduce band determined personal weakness criteria. This experience strengthened my belief in the value of the action research as a whole. I believed that this was a defining period for me also, in my attitude to risk taking. I believed that the following were important:

- A need for greater risk taking (for example, emphasising the importance of learning needs for individuals by increasing the focus on their band determined personal weaknesses).

- Evaluating my own practice and testing the stability of the emerging model (for example, making a decision to continue exploring the process features of phase two, as described in Section 3.5 of Chapter Three).

2004-2005

For the purpose of consolidating what I already understood, I felt that this concluding year should involve repeating aspects of the rehearsing and performance cycles. Nevertheless, as each cohort of popular music students has their own particular identities, I thought it dangerous to assume that peer assessment activities, to which previous students responded positively, would necessarily yield the same response this year. I felt that the following points were important for me in this final year:

- Increasing confidence in my practice (for example, it was with a sense of personal satisfaction that the activities which were repeated from the previous year did not appear to produce any particularly unexpected response).
- Learning about the importance of balance between risk taking and providing support for learning. For example, I accepted that the focus on three band determined personal weakness criteria only, for the purpose of emphasising their importance to individuals, might cause volatility in marking. Losing the smoothing effect, often apparent when employing a large number of criteria, was a risk that I was prepared to take, if this provoked individuals into improving their personal weaknesses.

Finally, at the time of writing (2008), the intra-peer assessment process model continues as current practice in providing the learning, teaching and assessment basis for Performance Management (now delivered at the University Centre Barnsley, a part of the University of Huddersfield).

7.6 Summary of the contribution to original knowledge

The research proposes a new process model of intra-peer assessment in popular music group rehearsing and performance that supports students' learning through the development of their personal attributes. Distinctive to this model is a graduated sequence of activities that identifies for each individual, a number of attributes, considered by themselves and by their peers, as necessitating development or improvement. It recommends a process (*band determined personal weaknesses*) whereby students acquire knowledge about themselves and attributes that others perceive in them as requiring development or improvement. The model, and the process of *band determined personal weaknesses* belonging to it, may extend our knowledge and understanding of peer learning among groups. The research might add to our understanding of how and why knowledge is constructed in this context. The enquiry as a whole found that, as a consequence of being their involved in this process, students may have developed an awareness and confidence about themselves, and acquired a particular moral knowledge of trust and honesty that related to interacting with their peers.

In addition, the thesis also offers a contribution to educational methodology in the development of a grounded theory approach to the application and analysis of quantitative peer assessment data.

The enquiry has been successful in the study of popular music, peer assessment, and group work rehearsing and performance in that it has revealed more about these interrelated disciplines within this curriculum area than any previously published research. This thesis adds to our understanding of peer assessment in supporting undergraduates' learning and the use of personal attributes as intra-peer assessment criteria for achieving this. The model has potential applications for other areas of the Higher Education curriculum in which knowledge creation, arising from the development of students' personal attributes displayed within group work, is a key aim.

Abson, D. (1994) The effects of peer evaluation on the behaviour of undergraduate students working in tutorless groups. In Foot, H., Howe, C., Anderson, A., Tolmie, A. and Warden, D. (eds.) *Group and Interactive Learning.* Southampton and Boston: Computational Mechanics.

Anderson, J.B. and Freiberg, H.J. (1995) Using self-assessment as a reflective tool to enhance the student teaching experience. *Teacher Education Quarterly*, 22: 77–91.

Aspin, D. (1982) The place of music: A justification. *Journal of Aesthetic Education.* 16/1, 41-5.

Atkinson, P. and Hammersley, M. (1994) Ethnography and participant observation. In Denzin, N. and Lincoln, Y. (eds.) *Handbook of qualitative research,* 248–261.Thousand Oaks, CA: Sage.

Bales, R.F. (1970) *Personality and interpersonal behavior.* NY: Holt, Rinehart, and Winston.

Ballantyne, R., Hughes, K. and Mylonas, A. (2002) Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment & Evaluation in Higher Education,* 27/5: 427-441.

Bandura, A. (1977) *Social Learning Theory.* New York: General Learning Press.

Bandura, A. (1986) *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall, Inc.

Bangert, A. (1995) *Peer assessment: An instructional strategy for effectively implementing performance - based assessments.* (PhD Dissertation, University of South Dakota, 1995). Dissertation Abstracts International, 56, 3505.

Belbin, R. (1981) *Management Teams.* NY: John Wiley & Sons.

Bhalerao, A. and Ward, A. (2001) Towards electronically assisted peer assessment: a case study, *Association for Learning Technology journal (ALT-J)*, 9/1: 26-37.

Biggs, J. (2003) *Teaching for Quality Learning at University.* 2nd ed. Buckingham: SRHE/Open University Press.

Bilgin, A. and Fraser, D. (2007) Empowering Students to be the judges of their own performance through peer assessment. In Phillips B. and Weldon L. (eds.) *The Proceedings of the ISI / IASE Satellite on Assessing Student Learning in Statistics.* International Statistical Institute, The Netherlands: Voorburg.

Billington, H. (1997) Poster presentations and peer assessment: novel forms of evaluation and assessment. *Journal of Biological Education*, 31/3: 218-220.

Birenbaum, M. (1996) Assessment 2000: towards a pluralistic approach to assessment. In Birenbaum, M. and Dochy, J. (eds.) *Alternatives in assessment of achievements, learning processes and prior knowledge. Evaluation in education and human services.* Boston, MA: Kluwer Academic Publishers.

Blom, D. and Poole, K. (2004) Peer assessment of tertiary music performance: opportunities for understanding performance assessment and performing through experience and self-reflection. *British Journal of Music Education,* 21/1: 111-125.

Bloxham, S. and West, A. (2004) Understanding the Rules of the Game: Marking Peer Assessment as a Medium for Developing Students' Conceptions of Assessment. *Assessment and Evaluation in Higher Education,* 29/6: 721-733.

Boehm, R.F. and Gallavan, N.P. (2000) Adapting Cooperative Learning to an Introductory Analysis Class. *Journal of Engineering Education,* 89/4: 419–421.

Boud, D. (1986) *Implementing student self-assessment: HERDSA green guide no. 5.* Kensington, N.S.W.: Higher Education Research and Development Society of Australasia.

Boud, D. (1991) *Implementing Student-Self Assessment.* Campbelltown, New South Wales: Higher Education Research and Development Society of Australasia.

Boud, D. (2000) Sustainable Assessment; rethinking assessment for the learning society. *Studies in Continuing Education,* 22/1.

Boud, D. and Walker, D. (1993) Barriers to Reflection on Experience. In Boud, D., Cohen, R. and Walker, D., (Eds) *Using Experience for Learning.* SRHE, Buckingham.

Boud, D., Cohen, R. and Sampson, J. (1999) Peer learning and assessment. *Assessment & Evaluation in Higher Education* 24/4: 413-426.

Boud, D. and Falchikov, N. (2006) Aligning assessment with long-term learning. *Assessment and Evaluation in Higher Education,* 31/4: 399-413.

Bransford, J., Brown, A. and Cocking, R. (2000) *How people learn (expanded edition): Brain, mind, experience, and school.* Washington DC: National Academy Press.

Brindley, C. and Scoffield, S. (1998) Peer assessment in undergraduate programmes. *Teaching in Higher Education,* 3: 79-89.

Brown, G. (2000) Peer learning and assessment: do they enhance student learning? In Hunter, D. and Russ, M. (eds.) *Peer Learning in Music, Conference Proceedings.* University of Ulster.

Brown, S., (ed.) (1998) *Peer Assessment in Practice.* SEDA Paper 102. Birmingham: Birmingham, Staff and Educational Development Association (SEDA).

Brown, S. (1999) Institutional Strategies for Assessment. In Brown, S. and Glasner, A. *Assessment Matters in Higher Education.* Buckingham: Society for Research into Higher Education and Open University Press.

Brown, S. and Dove, P., (eds.) (1991) *Self and Peer Assessment.* SEDA Paper 63. Birmingham: SEDA Publications.

Brown, S. and Glasner, A. (1999) *Assessment Matters in Higher Education.* Buckingham: Society for Research into Higher Education and Open University Press.

Brown, S. and Knight, P. (1994) *Assessing Learners in Higher Education.* London: Kogan Page.

Bruner, J. (1986) *Actual Minds, Possible Worlds.* Cambridge, MA: Harvard University Press.

Bryan, C. (2001) *Assessing Group Practice.* Notes from Peer Assessment workshop, Goldsmiths College, 7 Nov 2001. http://www.lancs.ac.uk/palatine/AGP/index.htm (accessed 29th Feb 2008).

Bryan, C. (ed.) (2004) *Assessing Group Practice.* SEDA paper 114. Birmingham: SEDA Publications.

Bryan, C. (2006) Developing Group Learning Through Assessment. In Bryan, C. and Clegg, K. (eds) *Innovative Assessment in Higher Education.* London: Routledge, Taylor and Francis Group Ltd.

Bryan, C. and Clegg, K. (eds) (2006) *Innovative Assessment in Higher Education.* London: Routledge, Taylor and Francis Group Ltd.

Bryman, A. (2004) *Social Research Methods* (2nd edition). Oxford: Oxford University Press.

Bushell, G. (2006) Moderation of peer assessment in group projects. *Assessment & Evaluation in Higher Education*, 31/1: 91–108.

Butcher, C. (2004) *Group Assessment.* Paper presented at the Higher Education Teaching and Learning Conference, Barnsley College, July 1-2.

Carr, W. and Kemmis, S. (1986) *Becoming Critical: Education, Knowledge and Action Research.* Philadelphia: The Falmer Press

Catterall, M. (1995) Peer learning research in marketing. In S. Griffiths, K. Houston, & A. Lazenblatt (eds.), *Enhancing student learning through peer tutoring in higher education.* Coleraine: University of Ulster, 1995.

Challis, M. (1999) AMEE medical education guide no. 11 (revised): Portfolio-based learning and assessment in medical education. *Medical Teacher,* 4: 437-440.

Charmaz, K. (2000) Grounded Theory: Objectivist and Constructivist Methods. In Denzin, K. and Lincoln, Y. (eds.) *Handbook of Qualitative Research* (2nd edition). Thousand Oaks, Ca.: Sage.

Charmaz, K. (2006) *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis.* London: Sage.

Cheng, W. and Warren, M. (1997) Having second thoughts: Student perceptions before and after a peer assessment exercise. *Studies in Higher Education*, 22/2: 233-240.

Cheng, W. & Warren, M. (1999) Peer and Teacher Assessment of the Oral and Written Tasks of a Group Project. *Assessment & Evaluation in Higher Education,* 24/3: 301–314.

Cohen, L., Manion, L. and Morrison, K. (2000) *Research methods in education* (5th edition). London: Taylor & Francis.

Conway, R., Kember, D., Sivan, A. and Wu, M. (1993) Peer Assessment of an Individual's contribution to a Group Project. *Assessment and Evaluation in Higher Education* 18/1: 45-56.

Cowan, J. (1998) *On Becoming an Innovative University Teacher*. Buckingham: Society for Research into Higher Education, RHE & Open University Press.

Creswell, J. W. (1998) *Qualitative inquiry and research design: Choosing among five traditions.* Thousand Oaks, CA: Sage.

Daniel, R. (2004) Peer assessment in musical performance: the development, trial and evaluation of a methodology for the Australian tertiary environment. *British Journal of Music Education*, 21/1: 89-110.

Davies, P. (2000) Computerized peer assessment. *Innovations in Education and Training International*, 37/4: 346-355.

Davies, P. (2003a) Closing the Communications Loop on the Computerized Peer-Assessment of Essays. *Association for Learning Technology Journal* 11/1: 41-54.

Davies, P. (2003b) Peer-Assessment: No marks required just feedback? – Evaluating the Quality of Computerized Peer-Feedback compared with Computerized Peer-Marking. In Cook, J & McConnell, D (eds.) *Communities of Practice, Research Proceedings of the 10th Association for Learning Technology Conference.* 8-10 Sept 2003, The University of Sheffield and Sheffield Hallam University, UK.

Davis, R. and Pulman, M. (2001) Raising Standards in Performance. *British Journal of Music Education* 18: 251-259.

Den Berg, I., Admiraal, W. and Pilot, A. (2006) Peer Assessment in University Teaching: Evaluating Seven Course Designs. *Assessment and Evaluation in Higher Education* 31/1: 19-36.

Denzin, N. and Lincoln, Y. (eds.) (1994) *Handbook of Qualitative Research.* Thousand Oaks: Sage.

Dochy, F., Segers, M. and Sluijmans, D. (1999) The use of self-, peer and co-assessment in higher education: a review. *Studies in Higher Education*, 24/3: 331 - 350.

Doran, S., Durston, C., Fletcher, A. and Longmore, J. (2000) Assessing students in seminars: evaluation of current practice. In Booth, A. and Hyland, P. (eds.) *The Practice of University History Teaching.* Manchester, U.K.: Manchester University Press.

Douglas, T. (1976) *Groupwork Practice*. London: Routledge.

Drew, S. (2001) Student perceptions of what helps them learn and develop in education. *Teaching in Higher Education,* 6/3: 309-331.

Edwards, R. and Sutton, A. (1991) A practical approach to student-centred learning. *British Journal of Educational Technology.* 23/1: 4–20.

Ellis, A. (2001) *Teaching, Learning, & Assessment Together: The Reflective Classroom.* Larchmont, NY: Eye on Education, Inc.

Exley, K. and Dennick, R. (2004) *Small group teaching.* Abingdon: RoutledgeFalmer.

Falchikov, N. (1986) Product Comparison and Process Benefits of Collaborative Peer Group and Self Assessment. *Assessment and Evaluation in Higher Education* 11/2: 146-166.

Falchikov, N. (1994) Learning from Peer Feedback Marking: Student and Teacher Perspectives. In Foot, H. *et al.* (eds..) *Group and Interactive Learning.* Southampton and Boston: Computational Mechanics Publications.

Falchikov, N. (1995a) Peer Feedback Marking: Developing Peer Assessment. *Innovations in Education and Training International*, 32/2: 175-187.

Falchikov, N. (1995b) Improving Feedback to and from Students. In Knight, P. (ed) *Assessment for Learning in Higher Education.* London: Kogan Page and SEDA.

Falchikov, N. (1996) Improving learning through critical peer feedback and reflection. *Higher Education Research and Development,* 19: 214-218.

Falchikov, N. (1998) Involving Students in feedback and assessment. In Brown, S., (ed.) (1998) *Peer Assessment in Practice.* SEDA Paper 102. Birmingham: SEDA Publications.

Falchikov, N. (2005) *Improving assessment through student involvement: practical solutions for aiding learning in higher and further education.* London: Routledge.

Falchikov, N. and Magin, D. (1997) Detecting gender bias in peer marking of students' group process work. *Assessment and Evaluation in Higher Education*, 22/4: 385-396.

Fallows, S. and Chandramohan, B. (2001) Multiple Approaches to Assessment: reflections on use of tutor, peer and self-assessment. *Teaching in Higher Education, 6*/2: 229-246.

Fosnot, C. (ed.) (1996) *Constructivism: Theory, Perspectives and Practice.* New York: Teachers' College Press.

Freeman, M. (1995) Peer assessment by groups of group work. *Assessment and Evaluation in Higher Education*, 20: 289–300.

Gallagher, D. (2004) The importance of constructivism and constructivist pedagogy for disability studies in education. *Disability Studies Quarterly*, 24/2: 1-15.

Gatfield, T. (1999) Examining student satisfaction with group projects and peer assessment. *Assessment & Evaluation in Higher Education.* 24/4.

Geertz, C. (1973) *Interpretations of culture.* New York: HarperCollins.

Gergen, K. (1991) The Saturated Self. New York: Basic Books

Gergen, K. (1995) Social construction and the education process. In Steffe, P. and Gale, J. (Eds.) *Constructivism in education.* Hillsdale, NJ: Lawrence Erlbaum.

Gibbs, G. (1992) *Improving the quality of student learning.* Bristol: Technical and Educational Services.

Gibbs, G. (1999) Using Assessment Strategically to change the way students learn. In Brown, S. and Glasner, A. (eds.) *Assessment Matters in Higher Education* Buckingham: Society for Research into Higher Education and Open University Press.

Gibbs, G. (2006) Why assessment is changing. In Bryan, C. and Clegg, K. (eds.), *Innovative assessment in Higher Education*, London: Routledge.

Glaser, B. and Strauss, A. (1967) *The Discovery of Grounded Theory.* Chicago: Aldine

Glaserfield, E, Von. (1991) Knowing Without Metaphysics: Aspects of the Radical Constructivist Position. In Steier, F. (ed.) *Research and Reflexivity*: 12-29. London: Sage.

Glaserfield, E, Von. (1995) *Radical Constructivism: A Way of Knowing and Learning.* London & Washington: The Falmer Press.

Goldfinch, J. and Raeside, R. (1990) Development of a Peer Assessment Technique for Obtaining Individual Marks on a Group Project. *Assessment and Evaluation in Higher Education*, 15/3: 210-231.

Goldfinch, J. (1994) Further Developments in Peer Assessment of Group Projects. *Assessment and Evaluation in Higher Education*, 19/1: 29-35.

Goodman, N. (1978) *Ways of Worldmaking.* Sussex: The Harvester Press.

Goodman, N. and Elgin, C. (1988) *Reconceptions in Philosophy and Other Arts and Sciences.* Indianapolis: Hackett.

Gray, T. (1987) An exercise in improving the potential of exams for learning. *European Journal of Engineering Education,* 12/4: 311-323.

Guba, E. and Lincoln, Y. (1994) Competing paradigms in qualitative research. In Denzin, N. and Lincoln, Y. (eds.) *Handbook of Qualitative Research*, Thousand Oaks, Ca: Sage.

Gupta, M. (2004) Enhancing student performance through cooperative learning in physical sciences. *Assessment & Evaluation in Higher Education,* 29/1: 63 -73.

Habeshaw, S., Gibbs, G. and Habershaw, T. (1993) *Interesting Ways to Assess Your Students*. Melksham: Cromwell Press.

Hanrahan, S. and Isaacs, G. (2001) Assessing Self and Peer-assessment: the Students Views. *Higher Education Research and Development*, 20/1: 53-70.

Hartley, P. (1997) *Group communication.* London: Routledge.

Heathfield, M. (1999) Group-based Assessment: An evaluation of the Use of Assessed Tasks as a Method of Fostering Higher Quality Learning. In Brown, S. and Glasner, A. (1999) *Assessment Matters in Higher Education.* Buckingham: Society for Research into Higher Education and Open University Press.

Heylings,.A. and Stefani, L. (1997) Peer Assessment Feedback Marking in a Large Medical Anatomy Class. *Medical Education* 33: 281-286.

Heywood, J. (2000) *Assessment in Higher Education.* London: Jessica Kingsley Publishers Ltd.

Hughes, I. and Large, B. (1993) Staff and peer-group assessment of oral communication skills. *Studies in Higher Education*, 18: 379-385.

Hunter, D. (1999) Developing peer-learning programmes in music: group presentations and peer assessment. *British Journal of Music Education* 16: 51-63.

Hunter, D. (ed.) (2004) *How Am I Doing? Valuing and Rewarding Learning in Musical Performance in Higher Education.* University of Ulster.

Hunter, D. (2006) Assessing collaborative learning. *British Journal of Music Education* 23/1: 75-89.

Hunter, D. and Russ, M. (1996) Peer assessment in performance studies. *British Journal of Music Education* 13: 67-78.

Hunter, D. and Russ, M. (eds.) (2000) *Peer Learning in Music.* University of Ulster.

Jacques, D. (2000) Learning in Groups, 3rd Ed., London: Kogan-Page.

Johnson, D., Johnson, R. and Holubec, E. (1993) *Cooperation in the classroom.* Edina, MN: Interaction Books.

Johnston, L. and Miles, L. (2004) Assessing contributions to group assignments. *Assessment & Evaluation in Higher Education*, 29/6: 751-768.

Kemmis, S. and Wilkinson, M. (1998) Participatory action research and the study of practice. In Atweh, B., Kemmis, S. and Weeks, P. (eds) *Action Research in Practice: Partnerships for Social Justice in Education.* London and New York: Routledge.

Keppell M, Au, E., Ma, A., and Chan, C. (2006) Peer learning and learning-oriented assessment in technology-enhanced environments. *Assessment and Evaluation in Higher Education*, 31/4: 453-464.

Kleiman, P. (2004) *To collaborate or not to collaborate: that is the question?* Shared Learning Workshop. York St. Johns. http://www.palatine.ac.uk/events/viewdoc/188/ (access date 28th Feb 2008).

Knight, P. and Yorke, M. (2003) *Assessment, Learning and Employability.* Berkshire UK: Open University Press.

Kwan, K. and Leung, R. (1996) Tutor versus peer group assessment of student performance in a simulation training exercise. *Assessment and Evaluation in Higher Education*, 21: 205–214.

Lapham, A. and Webster, R. (1999) Peer Assessment of Undergraduate Seminar Presentations: Motivations, Reflection and Future Directions. In Brown, S. and Glasner, A. *Assessment Matters in Higher Education.* Buckingham: Society for Research into Higher Education and Open University Press.

Langan, A., Wheater, C., Shaw, E., Haines, J., Cullen, W., Boyle, J., Penney, D.,Oldekop, J., Ashcroft, C., Lockey, L. & Preziosi, P. (2005) Peer assessment of oral presentations: effects of student gender, university affiliation and participation in the development of assessment criteria. *Assessment and Evaluation in Higher Education,* 30/1: 21-34.

Lave, J. and Wenger, E. (1991) *Situated learning: Legitimate peripheral participation.* New York: Cambridge University Press.

LeCompte, M. and Preissle, J. (1993) *Ethnography and qualitative design in educational research* (2nd edition). San Diego, CA: Academic Press

Lejk, M., Wyvill, M. and Farrow, S. (1996) A survey of methods of deriving individual grades from group assessments. *Assessment & Evaluation in Higher Education*, 21: 267–280.

Lejk, M. and Wyvill, M. (2001a) Peer assessment of contributions to a group project: A comparison of holistic and category-based approaches. *Assessment and Evaluation in Higher Education*, 26/1: 61-72.

Lejk, M. and Wyvill, M. (2001b) The effect of the inclusion of self-assessment with peer assessment of contributions to a group project: a quantitative study of secret and agreed assessments. *Assessment and Evaluation in Higher Education,* 26/6: 551-561.

Lejk, M. and Wyvill, M. (2002) Peer Assessment of Contributions to a Group Project: student attitudes to holistic and category-based approaches. *Assessment and Evaluation in Higher Education* 27: 569-577.

Li, L. (2001) Some Refinements on Peer Assessments of Group Projects. *Assessment & Evaluation in Higher Education.* 26/1: 5–18.

Lincoln, Y. and Guba, E. (1985) *Naturalist Inquiry*: London: Sage.

Lloyd-Jones, R. and Allen, J. (1997) Group Work, Assessment and the Journey to Graduate Employment: a Survey of History Students at Sheffield Hallam University. *Skills Versus Scholarship in Arts and Humanities in Higher Education: Conference Proceedings.* Milton Keynes: The Open University.

Longhurst, N. and Norton, L. (1997) Self-assessment in coursework essays. *Studies in Educational Evaluation*, 23: 319–330.

Lopez-Real, F. and Chan, Y. (1999) Peer Assessment of a Group Project in a Primary Mathematics Education Course. *Assessment & Evaluation in Higher Education*, 24/1.

Magin, D., Andresen, L., Nightingale, P. and Boud, D. (1992) *Strategies for Assessing Students*. Kensington: Professional Development Centre.

Magin, D. (1993) *Evaluating Peer Marking Procedures for Assessing Communication and Group Process Skills: Two Case Studies*. Paper presented at the Pacific Rim Symposium on Higher Education. Hilo, Hawaii, June 2-5.

Magin, D. (1996) Adjusting peer assessment scores to control for different rater standards. *Different Approaches: Theory and Practice in Higher Education.* Proceedings HERDSA Conference 1996. Perth, Western Australia, July 8-12.

Magin, D. (2001) A Novel Technique for Comparing the Reliability of Multiple Peer Assessments with that of Single Teacher Assessments of Group Process Work. *Assessment & Evaluation in Higher Education*, 26/2: 139-52.

McDowell, L. (1995) The impact of innovative assessment on student learning. *Innovations in Education and Training International*, 32: 302–313.

McDowell, L. and Sambell, K. (2005) *Peer Assessment and Autonomy: Student Perspectives*. Paper presented at the Professional Development Conference for Teachers in Higher Education, Staffordshire University, May 6th

McGourty, J., Dominick, P. and Reilly, R. (1998) *Incorporating Student Peer Review and Feedback into the Assessment Process*. Proceedings of the 1998 FIE Conference.

McNamara, M. and Deane, D. (1995) Self-assessment activities: toward language autonomy in language learning. *TESOL Journal,* 5/1: 17–21.

McNiff, J. (1988) *Action Research: Principles and Practice* (1st. Edition). London: Routledge.

McNiff, J., Lomax, P. and Whitehead, J. (1996) *You and Your Action Research Project* London: Routledge.

McNiff, J. and Whitehead, J. (2006) *Action Research: Living Theory* London: Sage Publications.

Mehrens, W., Popham, W. and Ryan, J. M. (1998) How to prepare students for performance assessments. *Educational Measurement: Issues and Practice*. 17/1:18-22.

Miles, M. and Huberman, A. (1994) *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.

Miller, P. (2003) The effect of coring criteria specificity on peer and self-assessment. *Assessment and Evaluation in Higher Education,* 28/4: 383-393.

Miller, L. and Ng, R. (1996) Autonomy in the classroom: peer assessment. In Pemberton, E. *et al.* (eds.) *Taking Control: Autonomy in Language Learning*. Hong Kong: Hong Kong University Press.

Mindham, C. (1998) Peer assessment: Report of a project involving group presentations and assessment by peers. In Brown, S., (ed.) (1998) *Peer Assessment in Practice.* SEDA Paper 102. Birmingham: SEDA Publications.

Morgan, W. and Wyatt-Smith, C. (2000) Im/proper accountability: towards a theory of critical literacy and assessment. *Assessment in Education: Principles, Policy & Practice,* 17: 123-142.

Mowl, G. and Pain, R. (1995) Using self and peer assessment to improve students' essay writing: a case study from geography. *Innovations in Education and Training International,* 32: 324-335.

Murtagh, L. (2007) Implementing a Critically Quasi-Ethnographic Approach. *The Qualitative Report,* 12 /2: 193-215.

Ney, J. (1989) Teaching English grammar using collaborative learning in university courses. *ERIC Document Reproduction Service.* ED 311463.

Nichol, D. and Milligan, C. (2006) Rethinking technology supported assessment practices in relation to seven principles of good feedback. In Bryan, C. and Clegg, K. (Eds) *Innovative assessment in Higher Education.* Oxford: Taylor and Francis Group Ltd, London.

Nicholson, T. and Ellis,G. (2000) Assessing group work to develop collaborative learning. In Booth, A. and Hyland, P. (eds.) *The Practice of University History Teaching.* Manchester, U.K.: Manchester University Press.

Norton, L. (2004) Using assessment criteria as learning criteria. *Assessment & Evaluation in Higher Education.* Vol 29/6: 687-702.

Neath, M. (1998) *The Development and Transfer of Undergraduate Group Work Skills.* (PhD thesis), Sheffield Hallam University.

Nystrand, M. (1997).*Opening Dialogue: Understanding the Dynamics of Language and Learning in the English Classroom.* New York: Teachers College Press.

O'Donnell, A. and Topping, K. (1998) Peers assessing peers: possibilities and problems. In: Topping, K. and Ehly, S. (eds.) *Peer-assisted learning.* Mahwah, NJ; London, Lawrence Erlbaum Associates.

Oldfield, K. and Macalpine, J. (1995) Peer and self-assessment at the tertiary Level - an experiential report. *Assessment and Evaluation in Higher Education,* 20: 125-132.

Oliver, R. and Omari, A. (1999) Replacing lectures with on-line learning: meeting the challenge. In Winn, J. (Ed.) *Responding to diversity: Proceedings of the 16th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education* (pp. 257-264). Brisbane: QUT.

Orsmond, P., Merry, S., and Reiling, K. (1996) The importance of marking criteria in the use of peer assessment. *Assessment & Evaluation in Higher Education,* 21: 239-250.

Orsmond, P., Merry, S., and Reiling, K. (1997) A study in self-assessment: Tutor and students' perceptions of performance criteria. *Assessment and Evaluation in Higher Education,* 22/4: 357 – 369.

Orsmond, P., Merry, S., and Reiling, K. (2000) The use of student derived marking criteria in peer and self-assessment. *Assessment and Evaluation in Higher Education,* 25/1: 23-38.

Orsmond, P., Merry, S., and Reiling, K. (2002) The use of examplars and formative feedback when using student derived marking criteria in peer and self-assessment. *Assessment and Evaluation in Higher Education,* 27/4: 309-323.

Pajares, F. (2002) *Overview of Social Cognitive Theory and of Self-Efficacy.* Emory University.http://www.des.emory.edu/mfp/eff.html (access date 29 Feb 2008).

Parsons, D. and Drew, S. (1996) Designing group project work to enhance learning: key elements. *Teaching in Higher Education,* 1: 65-80.

Pearce, A. (2000) Peer learning and assessment within a music curriculum designed for the promotion of personal attributes. In Hunter, D. and Russ, M. (eds.) *Peer Learning in Music, Conference Proceedings*, University of Ulster.

Peters, R. (1965) Education as initiation. In Archambault, R. (Ed.), *Philosophical analysis and education* (p. 110). London: Routledge Kegan Paul.

Piantinida, M. and Garman, B. (1999) *The Qualitative Dissertation.* Thousand Oaks: Sage Publications.

Pickford, R and Brown, S. (2006) *Assessing skills and practice.* Oxford, Routledge.

Pond, K., Rehan U., & Wade. W. (1995) Peer review: A precursor to peer assessment. *Innovations in Education and Training International,* 32/4: 314-323.

Pond, K. & Ul-Haq, R. (1997) Learning to access students using peer review. *Studies in Educational Research*, 23/4: 331-348.

Pope, N. (2001) An examination of the use of Peer Rating for Formative assessment in the Context of the Theory of Consumption Values. *Assessment and Evaluation in Higher Education* 26: 235-243.

Pope, N. (2005) The impact of stress in self- and peer assessment. *Assessment & Evaluation in Higher Education*, 30/1: 51-63.

Price, J. and Cutler, H. (1995) The development of skills through peer-assessment, in: Edwards, A. and Knight, P. (eds.) *Assessing Competence in Higher Education* Birmingham: SEDA.

Price, M. and O'Donovan, B. (2006) Clarifying Assessment Criteria and Standards. In Bryan, C. and Clegg, S., (eds.) *Innovative Assessment in Higher Education*, pp. 100-109, Routledge: New York.

Price, M., O'Donovan, B. and Rust, C. (2007) Putting a social-constructivist assessment process model into practice: Building the feedback loop into the assessment process through peer-feedback. *Innovations in Education and Teaching International*, 44 /2: 143-152.

Prins, F., Sluijsmans, D., Kirschner, P., & Strijbos, J. (2005) Formative peer assessment in a CSCL environment: A case study. *Assessment and Evaluation in Higher Education,* 30: 417-444.

Pulman, M. (2004) Peer Assessment in Popular Music. In Hunter, D. (ed.) *How Am I Doing? Valuing and Rewarding Learning in Musical Performance in Higher Education.* University of Ulster.

Purchase, H. (2000) Learning about interface design through peer assessment. *Assessment and Evaluation in Higher Education,* 27/4: 341-352.

Race, P. (1998) Practical pointers in peer assessment. In Brown, S., (ed.) *Peer Assessment in Practice.* SEDA Paper 102. Birmingham: SEDA.

Race, P. (1999) *Tips for Lecturers.* Kogan Page: London.

Rafiq, Y. and Fullerton, H. (1996) Peer assessment of group projects in civil engineering. *Assessment and Evaluation in Higher Education,* 21: 69-81.

Robson, C. (2002) *Real World Research,* 2nd ed. Oxford: Blackwell.

Rømer, T. (2002) Situated learning and assessment. *Assessment and Evaluation in Higher Education,* 27/3: 233-241.

Rudy D., Fejfar, M., Griffith, C. and Wilson, J. (2001) Self- and peer assessment in a first-year communication and interviewing course. *Evaluation & the Health Professions,* 24/4: 436-445.

Rust, C. (2007) Towards a scholarship of assessment. *Assessment and Evaluation in Higher Education,* 32/2: 229-237.

Rust, C., Price, M., and O'Donovan, B. (2003) Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment and Evaluation in Higher Education,* 28/2: 147-164.

Rust, C., Price, M. and O'Donovan, B. (2005) A social constructivist assessment process model: how the research literature shows us that this should be best practice. *Assessment & Evaluation in Higher Education,* 30/3: 231-240.

Sadler, D. (1987) Specifying and promulgating achievement standards. *Oxford Review of Education,* 13: 191-209.

Sambell, K., McDowell, L. and Sambell, A. (2006) Supporting diverse students: developing learner autonomy via assessment. In Bryan, C. & Clegg, K. (eds.) *Innovative assessment in Higher Education.* Oxford: Routledge.

Schön, D. (1987) *Educating the reflective practitioner: towards a new design for teaching and learning in the professions.* San Francisco, CA: Jossey-Bass.

Schwandt, T. (1994) Constructivist, interpretivist approaches to human inquiry. In Denzin, N. and Lincoln, Y. (eds.) *Handbook of Qualitative Research.* Thousand Oaks: Sage.

Searby, M. and Ewers, T. (1996) Peer assessing composition in higher education. *British Journal of Music Education* 13: 155-163.

Searby, M. and Ewers, T. (1997a) An Evaluation of the Use of Peer Assessment in Higher Education: a case study in the School of Music, Kingston University. *Assessment and Evaluation in Higher Education* 22: 371-383.

Searby, M. and Ewers, T. (1997b) Peer Assessment in Music at Kingston University. *The New Academic* 6: 5-7.

Sharp, S. (2006) Deriving individual student marks a tutor's assessment of group work. *Assessment and Evaluation in Higher Education*, 31/3: 329-343.

Shortt, K. (2001) *Staff and student benefits of increasing control over learning.* Liverpool: John Moores University. http://www.livjm.ac.uk/lid/lid_docs/ISSUE_02. (access date: 29th Feb 2008).

Siegel, S. and Castellan, N. (1988) *Nonparametric Statistics for the Behavioural Sciences.* New York: McGraw-Hill

Silverman, D. (1993) *Interpreting Qualitative Data.* London: Sage Publications.

Sitthiworachart, J. and Joy, M. (2004) *The Evaluation of Students' Marking in Web-Based Peer Assessment of Learning Computer Programming.* Paper presented at the International Conference on Computer in Education, Melbourne, Australia, Nov 30 - Dec. 3, 2004.

Sivan, A. (2000) The implementation of peer assessment: An action research approach. *Assessment in Education: Principles, Policy and Practice* 7/2: 193-213.

Skidmore, D. and Gallagher, D. (2005) *A Dialogical Pedagogy for Inclusive Education* Paper presented at the Inclusive and Supportive Education Congress, 1st - 4th August 2005. Glasgow.

Sluijmans, D., Dochy, F. and Moerkerke, G. (1998) *The use of self-, peer- and co-assessment in higher education : a review of the literature.* Educational Technology Expertise Centre. Heerlen: Open University of the Netherlands.

Sluijmans, D. and van Merriënboer, J. (2000) *A peer assessment model.* Educational Technology Expertise Centre. Heerlen: Open University of the Netherlands.

Sluijsmans, D., Moerkerke, G., Dochy, F. and van Merriënboer, J. (2001) Peer assessment in problem-based learning. *Studies in Educational Evaluation,* 27/2: 153-173.

Sluijsmans, D., Brand-Gruwel, S. and vanMerrienboer, J. (2002) Peer assessment training in teacher education: effects on performance and perceptions. *Assessment and Evaluation in Higher Education,* 27/5: 443-454.

Smith D., Scott E. and van der Merwe, N. (2007) Peer Assessment: A Complementary Instrument to Recognise Individual Contributions in IS Student Group Projects.

Electronic Journal of Information Systems Evaluation. Reading: Academic Conferences Ltd. 10/1: 61-70.

Smith, H., Cooper, A. and Lancaster, L. (2002) Improving the Quality of Undergraduate Peer Assessment: A Case for Student and Staff Development. *Innovations in Education and Teaching International: IETI,* 39/1, 71-87.

Somerville, H. (1993) Issues in Assessment, Enterprise and Higher Education: The Case for Self, Peer and Collaborative Assessment. *Assessment and Evaluation in Higher Education*, 18/3: 221-233.

Stahl, R. (1994) *Cooperative Learning In Social Studies: A Handbook For Teachers*. Menlo Park, CA: Addison-Wesley.

Stahl, R. and VanSickle, R. (eds.) (1992) *Cooperative Learning In The Social Studies Classroom: An Invitation To Social Study.* Washington, DC: National Council for the Social Studies.

Stefani, L. (1994) Peer, Self and Tutor Assessment: Relative Reliabilities. In Barnett, R. (ed.) *Studies in Higher Education*, 19/1: 69-75.

Stefani, L. (1998) Assessment in Partnership with Learners. *Assessment and Evaluation in Higher Education*, 23/4: 339-350

Strachan, I. and Wilcox, S. (1996) Peer and self-assessment of group work: developing an effective response to increased enrollment in a third-year course in microclimatology. *Journal of Geography in Higher Education*, 20: 343–353.

Strauss, A. and Corbin, J. (1998) *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (2nd Edition). London: Sage.

Sullivan, M., Hitchcock, M. and Dunnington, G. (1999) Peer and self-assessment during problem-based tutorials. *American Journal of Surgery*, 177: 266-269.

Swanson D., Case S. and Van der Vleuten, C. (1991) Strategies for student assessment. In (eds.) Boud, D. & Felletti, G. *The Challenge of Problem-based Learning.* London. Kogan Page. pp 260-274.

Topping, K. (1998) Peer assessment between students in colleges and universities. *Review of Educational Research* 68/3: 249-276.

Topping, K., Smith, E., Swanson, I. and Elliot, A. (2000) Formative peer assessment of academic writing between postgraduate students. *Assessment and Evaluation in Higher Education*, 25 /2:149-169.

Trevitt, A. and Pettigrove, M. (1995) *Towards autonomous criterion-referenced assessment and self-assessment: a case study.* Paper presented at the 2nd European electronic conference on assessment and evaluation: recent and future developments, Feb 6-10. European Association for Research into Learning and Instruction.

Tsai, C-C., Lin, S. and Yhan, S.-M. (2002) Developing science activities through a networked peer assessment system. *Computers & Education,* 38: 241-252.

Tuckman, B. (1965) Developmental Sequence in Small Groups. *Psychological Bulletin* 63: 384-399.

Tuckman, B. and Jensen, M. (1977) Stages of Small Group Development. *Group and Organizational Studies* 2: 419-427.

Vigotsky, L. (1976) *Mind in Society*. Cambridge, USA: Harvard University Press.

Watts, M. and Ebbutt, D. (1987) More than the sum of the parts: research methods in group interviewing. *British Educational Research Journal* 13/1:25-34.

Wells, G. (1999) *Dialogic Inquiry: Towards a Sociocultural Practice and Theory of Education*. Cambridge: Cambridge University Press.

Williams, E. (1992) Student attitudes towards approaches to learning and assessment. *Assessment and Evaluation in Higher Education*, 17: 45–58.

Wittgenstein L. (1953) *Philosophical Investigations.* (Translated by Anscombe, G.). Oxford: Blackwell.

Young, T. (1991) *Symbolic interactional theory and non-linear dynamics.* Transforming Sociology Series, 149. Michigan: The Red Feather Institute. http://critcrim.org/redfeather/chaos/ab-sit.htm(accessed 29th Feb 2008).

Zakrzewski, S. and Bull, J. (1998) The mass implementation and evaluation of computer-based assessments. *Assessment & Evaluation in Higher Education*, 23/2: 141-152.

Appendices

Appendix 1: Aims and learning outcomes of the modules

Appendix 2: Project brief given to students for a rehearsing and performing cycle

Appendix 3: Process used for calculating individual marks

Appendix 4: Framework of interview questions

Appendix 5: Interview analysis - transcript

Appendix 6: Interview analysis - data analysis procedures

Appendix 7: Example of an interview informing interventions in the cycles

Appendix 8: Quantitative assessment data analysis – examples of open coding, axial coding and selective coding

Appendix 9: Example of a band determined personal weaknesses agreement form

Appendix 1: Aims and learning outcomes of the modules

The following page contains extracts from the documentation for *Performance Management* module of the BA (Hons) Popular Music Studies programme validated by the University of Sheffield. The extract refers to the aims and learning outcomes that were specified for the module. The Performance Studies module shares the same learning outcomes. Some of the wording of the module aims differs slightly.

Barnsley College of Higher Education
BA (Hons) Popular Music Studies

Performance Management Module

This module aims to provide a basis to study the techniques involved in managing a variety of musical performances in diverse situations by focusing upon the development of performance skills, management abilities and collaborative groupwork.

Summary of Outcomes:
To achieve this unit a student must:
1. Perform the role of the musician in a group situation
2. Work effectively within a variety of musical contexts
3. Effectively use music technology in a performance where appropriate
4. Implement a prepared rehearsal schedule
5. Manage and organise public performances of the group
6. Critically evaluate the group's management of rehearsals and performances and the performance of others

Performance Studies Module

This module aims to provide a basis to study the techniques involved in managing a variety of musical performances in diverse situations by focusing upon the development of performance skills, management abilities and musical interactivity.

Summary of Outcomes:
To achieve this unit a student must:
1. Perform the role of the musician in a group situation
2. Work effectively within a variety of musical contexts
3. Effectively use music technology in a performance where appropriate
4. Implement a prepared rehearsal schedule
5. Manage and organise public performances of the group

Key skills: Managerial, leadership, organisational, interpersonal, decision making and problem solving, planning, applying technology.

Appendix 2: Project brief issued to the students

The following example is the project brief that was given to the *Performance Management* students for rehearsing and performing cycle 11 assignment: *Christmas Party.*

Project brief issued for the Christmas Party assignment, December 2002

BA1 Popular Music Studies	Assessment Criteria
Module Title Performance Management	As follows:
Project Title Decade tribute **Project Brief** You are to organise and provide music for your '*Christmas Party*', which is provisionally booked for the evenings of 17th, and 18th December at *The Assembly*, Barnsley. You should work within the group in which you have been placed. Each band should prepare a programme of 10-15 minutes performing time. You should plan your rehearsing time in agreement with the group members and in liaison with your tutor. You may use covers or originals. You will need to manage the logistics involved – equipment, set-ups, sound-checks and liase with the sound engineer. As rehearsing time and space is limited during the class sessions, your group may need to book rehearsal slots additionally at other times here in Honeywell, or elsewhere. Organising group rehearsals is a key element towards achieving success in the Performance Management module.	Students will be peer assessed as bands by BA3 Performance Techniques student panels using the criteria which you have generated: Tightness Technical control Dynamics and tuning Continuity Contrasting set Audience communication (visuals) Stage presence Composure/confidence Each band member will also be peer assessed on their contribution to the rehearsing based on: Three agreed *Group Attributes* Your three Individual *Attributes* Your final mark will be derived from: The mark awarded to your band factored together with the mark representing your contribution to the rehearsing, as peer assessed by your band members.
Assignment date: Nov 2002 Submission date: Dec 2002	Return date: Jan 2003 Tutors: MP/HD

Appendix 3: Process used for calculating individual marks

The following two pages provide illustrations of the method used for intra-peer marking and for determining the final mark awarded to each individual. The final mark for an individual student is calculated by combining the performance mark that was awarded to the band as a whole, with the intra-peer mark awarded for that band member. The marking scheme was explained to the students at the outset of the assignment and the calculations also appeared on the feedback mark sheets, an example of which can be found in this section of the Appendix.

The following is an example of an intra-peer assessment form. SC completed this for the *Christmas Party* assignment of 2004.

Christmas Party 2004 Band C	INTRA-PEER ASSESSMENTS (Personal weaknesses that you wish to improve in rehearsing) X = not applicable to you								
1 = negative/poor 2 = below average 3 = neutral/average 4 = above average 5 = positive/excellent Self assessment does not contribute to your mark	Punctual/timekeeping (Late/time wasting)	Willingness to help others (Not helping others)	Enthusiasm (Lacking enthusiasm)	Team member (Taking too much control)	Confident (Lacking in confidence)	Learning lyrics straight away (Leaving things until last minute)	Listening to others (Disregarding others)	Contributing to ideas (Holding back)	Being patient (Getting stressed)
HA	X	X	X	X	*3/5*	*4/5*	X	X	3/5
OM	3/5	X	X	X	3/5	X	X	2/5	X
SC	3/5	X	X	X	X	X	3/5	4/5	X
TP	X	5/5	*4/5*	*3/5*	X	X	X	X	X

The following is an example of a totalled intra-peer assessment form using three Group Attributes and three Individual Attributes (two *personal strengths* and one *personal weakness*).

Christmas Party Dec 2002 Band F	INTRA-PEER ASSESSMENTS											
	Group Attributes			Individual Attributes (Self chosen 2 personal strengths and 1 personal weakness) n/a=not applicable								
1-5 scale 1= negative poor 3= neutral average 5= positive excellent Personal weakness attributes appear in **Bold**	Reliability (Unreliability)	Creative input (No creative input)	Commitment (Poor commitment)	Performing Confident (Nervous)	Making decisions (Letting others do the deciding)	Excellent cooperative skills (Poor cooperation)	Excellent performing ability (Poor performing ability)	Remembering equipment (Not remembering equipment)	Organisation skills (Lack of)	Leadership (Poor leadership)	Excellent interpersonal skills (Poor interpersonal skills)	Punctuality (Not punctual)
SS	21/25	18/25	22/25	N/a	N/a	N/a	N/a	**16/20**	N/a	15/20	N/a	12/20
QB	13/25	10/25	12/25	11/20	9/20	N/a	N/a	N/a	N/a	N/a	N/a	**6/20**
LP	11/25	17/25	14/25	N/a	N/a	14/20	16/20	N/a	N/a	N/a	**13/20**	N/a
MJ	12/25	19/25	14/25	**14/20**	N/a	N/a	N/a	N/a	N/a	15/20	16/20	N/a
CR	23/25	18/25	24/25	**12/20**	N/a	N/a	N/a	N/a	19/20	N/a	N/a	18/20

The following is an example of the marking feedback that was given to each student (this was the written feedback to LA for the Christmas Party assignment of 2003). It presents the intrapeer assessment marking feedback and explains how the final individual mark is calculated.

LA Christmas Party 2003	Your Group Attributes			Your Individual Attributes (personal strengths and weaknesses; n/a=not applicable to you)			
Use 1-5 scale 1= poor/negative 3= average/neutral 5= excellent/positive Self assessment does not contribute to your mark	Attendance (Non-attendance)	Responsibility for learning parts (Not bothered)	Enjoyment/enthusiasm (Boredom/no enthusiasm)	Musical versatility (Non-music versatility)	**Taking constructive criticism (Not taking constructive criticism)**	Good playing (Bad playing)	Total
LA	13/15	8/15	10/15	12/15	**10/15**	10/15	63/90
Self assessment	5/5	3/5	3/5	3/5	**3/5**	5/5	

The Individual Attribute that you considered as a personal weakness (and one that you would like to improve) is printed in **Bold.**

BAND MARK AWARDED BY PEER PANEL=62%
YOUR INDIVIDUAL PEER ASSESSMENT TOTAL AWARDED BY YOUR BAND MEMBERS = 63/90
Average of Individual Peer Assessment Marks in your group: 71.6/90 (=100%)
Your Mark expressed as a % of the Average =63 divided by 71.6 (and multiplied by 100) = 88%
Your Final Mark = Band Mark (62%) multiplied by your individual percentage 88% and divided by 100.

YOUR FINAL MARK FOR THE CHRISTMAS PARTY REHEARSALS AND PERFORMANCE IS 55%

Appendix 4: Framework of interview questions

The following two pages provide an example of a semi-structured framework of questions used in interviews. As explained in Chapter Three interviews began with open, non-judgemental questions in order to encourage unanticipated responses to emerge (for example Question 1). The interview then moved to an exploration of peer assessment topics following the various responses given by the interviewees and guided by appropriate focused questions. It concluded with other open-ended questions in order to encourage interviewees to comment on areas that they thought had not been explored in their interview (for example Question 10). The format was flexible allowing additional questions to be introduced or students' responses to be explored more deeply.

Semi-structured interview: example of framework of questions

Question 1 (open question inviting responses to the intra-peer assessment experience):
You will remember assessing each of your band members, including yourself, as part of your overall grade for each performance last year. Tell me what you think or feel about the idea of assessing and awarding marks on the basis of individuals' rehearsing contribution based upon an agreed set of personal attributes.

Question 2 (focused questioning involving assessing band members)
Tell me what you think or feel about assessing each band member's contribution to rehearsing using personal attributes, knowing that these marks will actually count towards a student's individual grade?

Question 3 (focused questioning involving intra-peer assessment feedback and marks):
For the Venues & Audiences assignment last year - the schools gig – each band member was also asked after rehearsing for three weeks, to complete a rehearsal assessment that was 'advisory only'. This was to enable you to see how your band members thought that you were doing, so far up to that stage. The purpose being that of encouraging any student to try to improve upon any attribute in the remaining rehearsals before the next assessment where the marks would actually count. Here is the form for you to remind you. Tell me what you think or feel about this idea of using an advisory or formative assessment during rehearsing, followed by the formal assessment afterwards?

Question 4 (focused questioning about the timing of intra-peer assessment):
For the Christmas Party and Schools gigs, you assessed your band members' contributions after knowing what the overall band performance mark was. For the Seventies gig you assessed each member's contributions before knowing what the overall performance mark was for your band. Tell me whether knowing your performance mark in advance – be it disappointing, or unexpectedly good – influenced how you might have assessed your fellow band members?

Question 5 (focused questioning involving intra-peer assessment feedback and marks):
Do you think that a poor band performance mark might lead to students blaming certain band members excessively in terms of subsequently awarding low contribution marks, or, alternatively, given an unexpectedly good band performance mark, might band members overlook or forgive a band member who really deserved to be given a poor contribution to rehearsing mark?

Question 6 (focused questioning involving process of awarding marks in private and in collaborative settings):
You may remember that sometimes you were given the opportunity to assess each other's contribution to rehearsing together, collectively, with your band members. At other times you were instructed to assess each other privately, in confidence. Tell me what you think or feel about the effects of assessing in both of these ways?

Question 7 (focused questioning involving assessing band members):
I would like to remind you of how your various band members assessed your own contribution to rehearsing in terms of attributes that you stated. Remember that for both the schools gig and Seventies gig your individual attributes were those that you

considered were your personal 'weaknesses', or areas where you felt you wished to improve upon. (Give assessment profile to student). Looking at these assessments, over the year, your band members seem to think that you have improved in ...
Stayed about the same in... and declined in... Think carefully and take your time, and tell me honestly whether you think or feel that this is about right.

Question 8 (focused question about personal attribute criteria):
In using personal attributes to assess your fellow band members' contribution to rehearsing, tell me what effect do you think or feel that this had upon you?

Question 9 (open question about the intra-peer assessment experience):
Tell me what you have or have not learned about yourself and others in thinking about and assessing using personal attributes such as these?

Question 10 (open question inviting responses to issues that had not been explored):
Do you have any other comments to make from your experience of peer assessing your band members?

Appendix 5: Interview transcript

The following four pages provide a sample of an interview transcript. Full transcripts do not appear for reasons of confidentiality. The interviews were documented in this format with the numbers at the side relating to the open coding category analysis.

Extract of interview transcript for NL (interview conducted on17/12/2004).

MP Friday 17th December 2004, peer assessment interview; this is NL.
Thanks very much NL for attending the interview, I'm just confirming that it's ok to record?
NL: Absolutely
MP Last year you will remember rehearsing and at most times at some stage assessing your band members at the end of the assignments.
NL Yes
MP You will remember putting together various lists of personal qualities on which you assessed everybody and they assessed you, in the band. Just tell me how do you feel about the idea of assessing rehearsing in terms of qualities, personal attributes, rehearsing skills and so on – in the way that we did it?
NL I think it's a very good idea; I think it helps to pick out the people who put in hard work....
1 I think it helps to bring to attention anyone who has any slight faults, or needs to brush up upon anything; it helps to do that. It gives a good constructive criticism to each other without sounding picky, without being nasty. I think its good that you can choose your
2 different aspects of things, you know, and what, and actually choose what you think your strengths and weaknesses are. 'Cos it's not easy for people to pick their strengths
3 and weaknesses and, I think it really makes you assess yourself, and look at yourself and see what you're doing wrong, what you are doing right, what you could do better...
MP OK. Thanks. So, there were some occasions where you were asked individually, all of you, to name personal qualities, or, if you like 'weaknesses' – things that you feel you could do better at in a rehearsal situation and, at other times I actually asked your band members to come up with qualities as well, that they thought of you personally, because they rehearsed with you, they know you better than anybody else at that stage
NL Yes
MP ... to come up with things that they felt you might want to develop. Was that helpful... or was it difficult to cope with, you know, in terms of – it can get quite personal, I suppose in that situation, where other students are actually setting your attributes for you. How did you feel that went?
NL I think, if somebody else is willing and brave enough to come up with your faults,
4 then I think you have to listen to them, because they obviously think that there is an
5 issue, and something that could be made better or resolved. Maybe something that you
6 personally wouldn't look at for yourself. And I think that is a good idea. Any criticism is hard for anybody, but if it makes you take stock and look at what you are doing, then what's the harm in it?

MP Was there – there were occasions when you assessed – all of you - assessed each other before the actual performance (i.e. you assessed each other's rehearsing before the band actually performed – certainly before receiving your initial overall performance mark. There was another occasion where you were told what your marks were, then you assessed each other. Would knowing in advance, what your overall band performance mark is, would that actually influence you, do you think, or any others, when you were then subsequently asked to assess individuals in rehearsing?

7 NL Possibly. Because I would say – if somebody was to – well If I was to be completely honest – if I was to find out that somebody gave me, say, 2 out of 12 for something that I wasn't as good at, then I might be tempted to mark somebody slightly lower, because I was upset about that. That's not necessarily how everybody would react. But I think its better not to find out until afterwards. I really do. You couldn't be bias anybody.

MP Yes. So you feel that when the band received the performance mark from a gig that was disappointing, or, on the other hand, surprisingly good, you think that actually might cause people to apportion blame, and so on?

8 NL Yes, Absolutely, but also in saying that I do know a particular thing that happened – I'm not going to name names as such, but last year in somebody's band they got marked lower and they thought they'd put a lot of work in. - and what they said that was next
9 time they do that I'm going to deliberately mark the other people lower because they deliberately gave low marks this time. It's just kind of tit for tat. But – people, you know, find it hard not to retaliate with this kind of thing.

MP Yeah, yeah. I'm aware that is human nature - that is inevitable to some extent. In that case then, are there any ways in which we could improve assessments along these lines? How could we do it better – given that there will always be – inevitably - one or two instances where there is that tit –for- tat situation?

10 NL Well I mean as it is now it's fine, because, you don't see who's marked you, individually, you just see the total, which is good, the only way around would be not to give those marks out till say, all the performances have finished, but in reality that won't work because people will want to know how they've done. I think it works well this way. I really do. If I can see a major flaw in it – not even a major flaw – what I think that if somebody, doesn't actually – in a band if a couple of people don't jell – just on a personal basis – you've got to look ahead, you've got to mark in a professional way that when you're doing this, each student has to look at it in a level headed kind of manner and take your personal feelings to one side and look at the person's talent rather than personality, and what you think of them outside the group.

MP Yes. I think you are right; it takes some maturity perhaps to separate personal issues from the professional job of rehearsing and the kinds of qualities...

NL And I also think that if in a rehearsal if somebody – 'could you please do that
11 differently' – if, you know, if things had been brought up, then somebody might think, 'right, well, they've been not tolerant because of that', rather than think it as, you know, just a piece of constructive criticism. That's it - the only downside I can to it. Other than that I see it's an exceptional way of getting feedback, because who's better to give you response than the peers that you are working with.

MP That's right, you spend most time with them, a lot more time rehearsing than in the performance ...

NL Yeah

MP ...and, as you say, it's right that your fellow band members are in the best place. There were times when I asked you – or rather I invited bands if they wished to, whether they wanted to assess privately, independently - in secret if you like - or, whether you wanted to discuss marks as a group and, you know sensibly, fairly and with maturity, apportion marks. Some preferred marking in private. What are your views on the actual marking?

NL Well I think, if you are doing a performance, everybody's there to see it, I think everybody should know what the reactions are. It's the same if you go to see a major concert, you read the reviews in the paper, - you read the critics marks, and they are

open to everyone to view. I just find that everybody can be vocal and talk about it. If somebody disagrees then they can bring it up.
MP So, are you suggesting that the band can sit down – this is before the performance – at the end of the rehearsing, the last rehearsal, if you like, with an assessment 'grid' in front and discuss and decide fairly and equitably where the marks really ought to be divided?
NL They could, but in a realistic world that wouldn't happen, I suppose. I got confused there, I thought you meant the performance assessment – sorry,
MP Right, sorry
NL No, I think if you are going to mark somebody like that – (I'm sorry but I
12 completely got the wrong end) - it should be a private thing, but I do think you should keep personal feelings out. No: I don't think it can be discussed amongst others. If you want to give your mark to someone it would bias their views, it can blur the edges, it's not a good idea.
MP We've always, on giving feedback, everybody's marks have been totalled and summed together so there's no way that an individual will know how another individual has marked them on that. Are you happy with that, or do you think we ought to have more of individuals' feedback?
NL No. I really don't. Because as I say, criticism can hurt, and if it comes from somebody that you don't expect it from, it can hurt even more. And I don't think it's such a good idea that it comes back. It's better the way it's done now. Personally, that's how I feel.
MP What have you learned using these personal qualities, personal attributes, etc. What have you learned about yourself, really, in the responses and marks that have been given to you over the three assignments? We started off with the Christmas Party assignment – that was when you were put together in groups by myself for that. The Venues & Audiences – the 'schools gig, – that was the next one, and you chose who you wanted to work with, and, the same for the final gig, which was the 'Eighties' gig. The 'Schools gig' and the 'Eighties gig' you continued those 'qualities' for both of those gigs– they were unchanged.
NL Yes
MP If you just have a look at, just to remind yourself, the Christmas –you were in two bands...
NL Yes
MP Then the schools gig and the eighties gig. What have you generally learned – if anything! – about yourself, in the responses that have been given to you?
13 NL Definitely to be more tolerant. Definitely. I'm not a very tolerant person; I get upset when people don't turn up for rehearsals, when they've been organised but don't have the trouble to come in. And then, in a later rehearsal, when things don't go right, I very, very was quick to judge and to say: "if you'd been here, we wouldn't be in this
14 position." I am, I have learned to become a lot more tolerant of people because that was one thing that I brought up myself that I thought was my personal weakness, and which was that someone else in my band brought up, to do with my tolerance. I've definitely learned to be more tolerant. I also, wanted to have a more positive outlook, which was one of the things I brought up myself, because I'm always sitting there saying 'oh, we're never going to do it, it's not going to happen...' But I'm beginning to be a more positive, and I think that's through everyone else in the band being more positive as well. Definitely helps.
MP Mm, good. Okay.

Looking at the range of marks there, I know it seems quite a while ago now, perhaps difficult to remember, looking at the marks would you say they are about right, or do you have issues with some.... tell me what your responses are...

NL Well, I'm a singer, that's what I do, I'm not a band player as such, not a lead guitar player, I actually find that with this, a lot of the times I have been completely in the wrong band. Myself, I am not musically happy, I don't feel safe letting myself go to sing how I want to sing because I'm worried that the people behind me aren't going to pull it off – and that sounds very big-headed and very bad but, I think the marks are right because the actual... it's not always jelled as well as it had done. I thought when we did 'venues and audiences' we had a very, very tight band, it was very good. When we did the 'Eighties Night' it was just the three of us and we did three songs in one day because we'd been in a band or the rest of the band. Basically yes, in fact the eighties
15 night was a bit higher than I would have expected to have got - but yes, they're fair. Obviously I would have liked higher!

MP OK, I appreciate your comments and you've mentioned some very interesting things there.... Anything else that you'd like to say about assessing and evaluation of rehearsals in this way – just take a moment to consider – and also, say the process of doing this, from a tutor point of view. I've mentioned a couple of things where a tutor would have to decide, for example, whether to invite a band to collaborate in order to come up with fair marks; another 'process' would be when actually to do this? Would there be any mileage, for example, in doing this say half way through the duration of an assignment, - after three weeks or so – and then perhaps repeating it to see whether there is any difference...?

NL I think you need to know about it because you need to be thinking what you are going to talk about and what you are going to make as your weakness, your strength and other peoples' weakness and strengths. I think to spring it on to you for about 3 weeks into it, you maybe haven't looked at people and read them as well as you might have
16 done. I know about this, you're constantly looking and thinking what are your strengths and weaknesses. You know and its maybe not so obvious, you maybe see an underlying strength or weakness, which you wouldn't see if it was just sprung on you at the last minute – to say get this done by the next week.

MP Did you make a conscious effort, or decision, or whatever when knowing that you would be assessed at some stage by members, across those attributes; were you remembering this during rehearsals, were you consciously... or was it vaguely in the
17 background?

NL it was vaguely in the background, consciously thinking well " I'm going to be peer assessed here" so you are constantly trying to better yourself.... Yes, once you know that you're going to be assessed like that, it's constantly there in rehearsals.

Appendix 6: Data analysis procedures

The following pages provide examples and extracts of grounded theory that was adapted to the analysis of the interviews with *Performance Management* students.

Appendix 6a is an extract of the open coding category and thematic analysis of the interview that was conducted with NL, appearing above.

Appendix 6b is an extract from the axial coding analysis of the core category: *Selection of Personal Attributes*. Sub-categories identified within this core category are indicated in bold text.

Appendix 6a: Data analysis procedures

Interview analysis of NL: extract of open coding categories and themes

	Open coding categories and themes
1	P/a picks out people who put in hard work [identification of personal attributes]
2	Constructive criticism without being picky or being nasty [honesty, constructive feedback]
3	Good to be able to choose qualities [self awareness, self assessment, validation]
4	You have to listen and respect someone brave enough to tell you your faults [respect/trust]
5	Might come up with something you hadn't considered yourself [awareness of others]
6	If criticism makes you take stock then P/a is OK [impact of intrapeer feedback]
7	Intra-peer assessment without knowing performance mark [process encourages honesty]
8	Unfair marking can lead to retaliation [bias, tension]
9	Best not to see who has marked you individually [trust, honesty]
10	Need to mark in a professional manner, not on personal feelings [desire honest marking]
11	An exceptional way of getting feedback [validation, feedback]
12	Assessment in private, not collaborative [process encourages honesty]
13	Intra-peer assessment developed my tolerance [develops personal attributes]
14	Intra-peer assessment developed a more positive outlook in myself [develops personal attributes]
15	Intra-peer assessment rehearsing marks were fair [fairness, validation]
16	Formative p/a after 3 weeks is not enough time to consider weaknesses [questionable process]
17	Knowing one is to be p/a by band constantly there in the background [thinking, cognition]

Appendix 6b: Data analysis procedures

Extract of axial coding category: Selection of Personal Attributes. Axial coding sub-categories that were identified appear in **bold**.

	Core Category: Selection of Personal Attributes															
162	**Band members to choose attributes, know you better**			HA	DC	SA			WK	SH		EN				
163	**Perceptions of self-weaknesses are not what others perceive as weaknesses**			HA		SA	NL									
	Remaining coding of Selection of personal attributes:															
164	Band member choose because more honest identifying weaknesses					SA			WK							
165	Group to decide what is important for attributes					SA			WK							
166	6 attributes is good and spreads attention over things					SA										
167	Individual members may deliberately choose a strength for a weakness		OD													
168	Good to be able to choose qualities						NL									
169	Good to concentrate on some of the areas of personal attributes							OM								HS
170	Band members choosing your weaknesses must have your agreement		OD													
171	Being inventive is bringing your own personality into the rehearsing											EN				
172	Don't agree with everything that's picked for me											EN				
173	Agree with attributes chosen by band	YL									TP					
174	Student picks an extra attribute chosen by individual, for feedback only	YL														
175	P/a leads to awareness of other people's attributes, good for a band situation				DC											
176	Attributes focused me better				DC											HS
177	Choose one attribute get it to a high level, then move to another				DC											
178	Deciding own criteria leads to what we are good at									SH						
179	Frustrated by non-inclusion of an attribute that should be marked down									SH						
180	Discuss attributes with band so they appreciate what you are trying to do								WK							
181	They're qualities that I would uphold all the time												EJ			
182	Versatility means marks for changing instruments and stuff like that													DC		
183	Leadership, means lead most of time but step back and let others also													DC		

Appendix 7: Example of an interview informing interventions in the cycles

The following page is an example of where an interview informed a number of interventions in the action research cycles that immediately followed. The interview was subsequently subjected to grounded theory analysis.

Appendix 7: Example of an interview extract, which informed subsequent interventions in the cycles.

Interview with SA 24th November 2003, following cycle 13.

Comments by SA informing subsequent interventions	Implied interventions	Interventions in subsequent cycles
SA: Possibly your band members could choose the attributes for you... I think it would be better, actually, if your band members, if they know you well enough, to come up with those things think you need to improve on, instead of you picking them yourself.	Using band determined personal weaknesses	15, 16, 18, 19
SA: With your group attributes, get in a group and decide what is important	Using group attributes	14,15, 16
SA: If you are assessing whilst in a group of people, you are more likely to favour people by not looking as though you are being stingy or harsh. And someone will say " are you really going to give me this for that?" and then it will seem like... and people will be genuinely offended, whereas if you did it privately... you are more likely to get a fair description of what you think happened... I would prefer to take it home and think about it, instead of having to do it in 10 minutes.	Conducting intrapeer assessment in private	14,15,17,18

Appendix 8: Quantitative assessment data analysis

The following pages indicate the extent of the extensive and 'thick' descriptive quantitative data that was collected and analysed during the course of the action research.

Appendix 8a provides an example (Venues & Audiences 2004, Band A) of an intra-peer assessment form in which the band members' marks are assembled prior to analysis. 75 of these forms were assembled. Appendix 8b is an example of the open coding process that was applied to the intra-peer assessment data (Venues & Audiences 2004, involving Bands A, B, C and D). 20 of these forms were assembled.

Appendix 8c is an example of the standard deviations analysis relating to an intra-peer assessment (Venues & Audiences 2004, Band A). 75 of these forms were assembled. Appendix 8d is an example of the open coding process that was applied to the standard deviation data (Venues & Audiences 2004, involving Bands A, B, C and D). 20 of these forms were assembled.

Appendix 8e is an example of the normalised difference coefficients analysis that was applied to an intra-peer assessment (Venues & Audiences 2004, Band A). The formula that was used to calculate normalised difference coefficients is also provided. Over 60 of these forms were assembled. Appendix 8f is an example of open coding process that was applied to the normalised difference coefficient data (Venues & Audiences 2004, involving Bands A, B, C and D). Around 20 of these forms were assembled.

Appendix 8g illustrates the axial coding process for standard deviation data (collected during the 2003-2004 academic year). Appendix 8h illustrates the axial coding analysis summary matrix for the Venues and Audiences assignments, 2004. Appendix 8i illustrates the selective coding of the analysis summary matrix for 2003-2004. Appendix 8j illustrates the selective coding and integration of the analysis summary matrix 2003-2004.

Appendix 8a: Example of an intra-peer assessment form in which the band members' marks are assembled prior to analysis.

March 2004 Venues & Audiences Band A	INTRA-PEER ASSESSMENTS										
	Individual Attributes (Personal 'weaknesses' in rehearsing as identified by band) x =not applicable									Group Attributes	Total
Self-assessments appear in **bold**	Contributing to ideas (Not contributing to ideas)	Musical Versatility (Unwilling to change to other instruments)	Punctual (Late)	Stage presence (Unprepared stage presence)	Adapting to different styles (Difficult adapting…)	Leadership skills (Allowing others to take leadership role)	Attendance (poor attendance)	Team member (not team member)	Enthusiasm (unenthusiastic)		
SK	43333 *13/20*	x	x	33334 *12/20*	34334 *13/20*	x	x	x	x	4544 *17/20*	55/80
RD	43543 *16/20*	x	23253 *12/20*	x	34544 *17/20*	x	x	x	x	4454 *17/20*	62/80
HM	44543 *17/20*	x	33533 *14/20*	x	x	33544 *15/20*	x	x	x	4454 *17/20*	63/80
DC	x	45443 *17/20*	35533 *16/20*	x	x	35543 *17/20*	x	x	x	4544 *17/20*	67/80
GM (did not self-assess)	x	x	x	x	x	x	3333 *12/20*	3333 *12/20*	3333 *12/20*	3333 *12/20*	36/60

Appendix 8b: Example of open coding of intra-peer assessments

	Coding of Intra-peer Assessments Venues & Audiences 2004	Bands and band members			
Coding Ref:	Open coding categories	Band A SK/RD/ HM/DC/GM	Band B MS/DS/SA GT/GR	Band C DT/LH/ GT/PA/G R	Band D HW/EJ/LM NL/LA/SH/J S
04vaia1	Group attributes assessed higher than personal weaknesses	4	4	2	5
04vaia2	Group attributes and personal weaknesses assessed equally	1	1	1	1
04vaia3	Group attributes assessed lower than personal weaknesses	0	0	2	1
04vaia4	High consist of agreement in assessing personal weaknesses (variance of 1)	4	7	13	8
04vaia5	Fair consist of agreement in assessing personal weaknesses (vari + or - 1)	7	6	2	10
04vaia6	Low consist of agreement in assessing personal weaknesses (vari >+ or - 1)	0	2	0	3
04vaia7	High consistency of agreement in assessing group attributes (variance of 1)	4	2	4	5
04vaia8	Fair consistency of agreement in assessing group attributes (vari of + or - 1)	1	3	1	2
04vaia9	Low consistency of agreement in assessing group attributes (vari >+ or - 1)	0	0	0	0
04vaia10	High consistency of agreement in assessing freerider (variance + or - 1)	4	3	3	3
04vaia11	Fair consistency of agreement in assessing freerider (variance >+ or - 1)	0	0	1	1
04vaia12	Low consistency of agreement in assessing freerider (variance 1)	0	1	0	0
04vaia13	High consistency of agreement in assessing shootingstar (variance + or - 1)	2	1	4	0
04vaia14	Fair consistency of agreement in assessing shootingstar (variance >+ or - 1)	2	3	0	4
04vaia15	Low consistency of agreement in assessing shootingstar (variance 1)	0	0	0	0
04vaia16	Freerider group attributes assessed lower than 3 personal weaknesses	Equal (GM)	No (MS)	No (GT)	Equal (LA)
04vaia17	Freerider group attributes assessed higher than 3 personal weaknesses	Equal (GM)	Yes (MS)	Yes (GT)	Equal (LA)
04vaia18	Shootingstar group attributes assessed lower than 3 personal weaknesses	No (DC)	Equal (GR)	No (PA)	No (LM)
04vaia19	Shootingstar group attributes assessed higher than 3 personal weaknesses	Yes (DC)	Equal (GR)	Yes (PA)	Yes (LM)

Appendix 8c: Example of standard deviations analysis of an intra-peer assessment

March 2004 Venues & aud Band A	INTRA-PEER ASSESSMENTS									
	Individual Attributes (Personal 'weaknesses' in rehearsing as identified by band) x =not applicable									Group Attributes (general contribution to rehearsing)
Mark range 1 -5	Contributing to ideas (Not contributing to ideas)	Musical Versatility (Unwilling to change to other instruments)	Punctual (Late)	Stage presence (Unprepared stage presence)	Adapting to different styles (Difficult adapting...)	Leadership skills (Allowing others to take leadership role)	Attendance (poor attendance)	Team member (not team member)	Enthusiasm (unenthusiastic)	
SK	0.5	x	x	0	0.5	x	x	x	x	0.5
RD	1	x	1.26	x	1	x	x	x	x	0.5
HM	0.58	x	1	x	x	0.96	x	x	x	0.5
DC	x	0.5	1.15	x	x	0.96	x	x	x	0.5
GM	x	x	x	x	x	x	0	0	0	0.0

Appendix 8d: Example of open coding of standard deviations

	Coding of Standard Deviations Venues & Audiences 2004				
Reference	Categories	Band A	Band B	Band C	Band D
	Band members	SK/RD/ HM/DC/ GM	MS/DS/ SA/GT/ GR	PA/DT/ GR/LH/ GT	HW/EJ/LM NL/LA/ SH/JS
04vasd01	Personal weaknesses exhibit small standard deviation (0.5 or smaller) overall	8	7	15	7
04vasd02	Personal weaknesses exhibit average standard deviation (0.6 – 1.2) overall	7	7	0	12
04vasd03	Personal weaknesses exhibit large standard deviation (1.3 or larger) overall	0	1	0	2
04vasd04	Group attributes exhibit small standard deviation (0.5 or smaller))	5	2	4	5
04vasd05	Group attributes exhibit average standard deviation (0.6 – 1.2)	0	3	1	2
04vasd06	Group attributes exhibit large standard deviation (1.3 or larger) overall	0	0	0	0
04vasd07	Personal weaknesses has smaller mean standard deviation than Group attribute	1	1	4	2
04vasd08	Personal weaknesses and Group attribute have same mean standard deviation	1	1	1	0
04vasd09	Personal weaknesses has larger mean standard deviation than Group attribute	3	2	0	5

Appendix 8e: Example of normalised difference coefficients analysis of an intra-peer assessment

March 2004 Venues & Audiences Band A	NORMALISED DIFFERENCE COEFFICIENTS Individual Attributes (personal weaknesses)										
	Contributing to ideas (Not contributing to ideas)	Musical Versatility (Unwilling to change to other instruments)	Punctual (Late)	Stage presence (Unprepared stage presence)	Adapting to different styles (Difficult adapting...)	Leadership skills (Allowing others to take leadership role)	Attendance (poor attendance)	Team member (not team member)	Enthusiasm (unenthusiastic)	Mean Student Coefficient	Band coefficient
SK	0.95	x	x	0.7	0.85	x	x	x	x	0.58	
RD	0.75	x	0.87	x	1.0	x	x	x	x	0.87	
HM	0.69	x	0.87	x	x	0.94	x	x	x	0.84	0.75
DC	x	0.69	0.75	x	x	0.69	x	x	x	0.71	
GM	x	x	x	x	x	x	N/a	N/a	N/a	N/a	

(As GM did not provide any self-assessments, it was not possible to perform normalised difference coefficients)

Normalised difference coefficients, or inverse measure of proximity normalised between 1 (=good/maximum) and 0 (=poor/minimum) based on 1-5 marking scale, comparing self-assessment with mean of the remainder of the band, where x = mean and y = self assessment.

$$\frac{x-y}{5-1} = \left\{ \frac{1x-y}{4} \right\}$$

Appendix 8f: Example of open coding of Normalised Difference Coefficients

	Coding of Normalised Difference Coefficients Venues & Audiences 2004				
Reference	Categories	Band A	Band B	Band C	Band D
	Band members	SK/RD/ HM/DC/ GM	MS/DS/ SA/GT/ GR	DT/LH/ GT/PA/ GR	HWEJ/LM NL/LA/ SH/JS
04vand01	Personal weaknesses exhibit a normalised difference coefficient of 0.90 or higher	3	4	6	5
04vand02	Personal weaknesses exhibit a normalised difference coefficient of 0.75 or higher	5	7	6	5
04vand03	Personal weaknesses exhibit a normalised difference coefficient of less than 0.75	4	4	3	5
04vand04	Group attributes exhibit a normalised difference coefficient of 0.9 or higher				
04vand05	Group attributes exhibit a normalised difference coefficient of 0.75 or higher				
04vand05	Group attributes exhibit a normalised difference coefficient of less than 0.75				
04vand07	Personal weaknesses have consistently smaller normalised difference coefficient than Group attribute				
04vand08	Personal weaknesses and Group attribute are similar				
04vand09	Personal weaknesses have consistently larger normalised difference coefficient than Group attribute				
04vand10	Mean student coefficient of 0.90 or higher	0	1	1	2
04vand11	Mean student coefficient of 0.75 or higher	2	3	4	2
04vand12	Mean student coefficient of less than 0.75	1	1	0	1

Appendix 8g: Example of axial coding of standard deviations

	Axial Coding of Standard Deviations Combined 2003-2004	
Reference	Categories	
04cond01	Personal weaknesses exhibit a small standard deviation (0.5 or smaller) overall	111
04cond02	Personal weaknesses exhibit an average standard deviation (0.6 – 1.2) overall	75
04cond03	Personal weaknesses exhibit a large standard deviation (1.3 or larger) overall	18
04cond04	Group attributes exhibit a small standard deviation (0.5 or smaller))	70
04cond05	Group attributes exhibit an average standard deviation (0.6 – 1.2)	29
04cond06	Group attributes exhibit a large standard deviation (1.3 or larger) overall	4
04cond07	Personal weaknesses smaller mean standard deviation than Group attribute	18
04cond08	Personal weaknesses and Group attribute have same mean standard deviation	5
04cond09	Personal weaknesses has larger mean standard deviation than Group attribute	29

Appendix 8h: Example of axial coding: analysis summary matrix

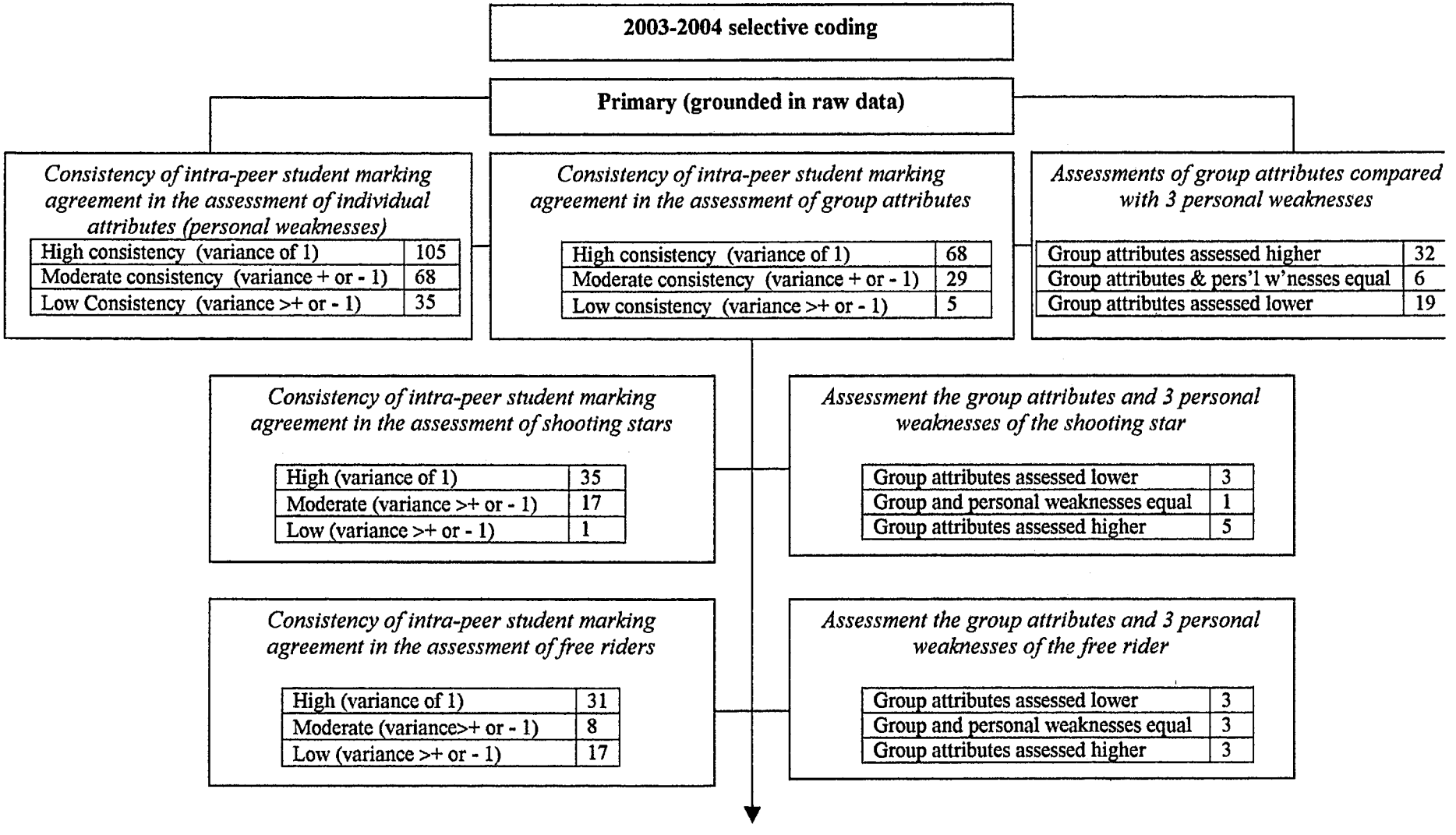
VENUES & AUDIENCES MARCH 2004 Analysis Summary Matrix										
		Individual Attributes				Process features		Group Attributes	Inter-peer Assessment	Individual Final Grade
Student (*= absent, or, did not provide self assessment mark)		Mean of combined Standard Dev of intra peer assessment (excl. self-assess)	Self Assessment Variance (raw marks)	Individual Normalised Difference Coefficient	Band Normalised Difference Coefficient	Private or collaborative assessment	Assessment conducted before or after receiving performance mark	Standard deviations of group attributes	Performance Assessment mark awarded by peer panel	Final combined mark for Venues & Audiences
BAND A	SK	0.33	+1.5	0.58	0.75	C	Before	0.5	62%	58%
	RD	1.09	-1.25	0.87	0.75	C	Before	0.5	62%	65%
	HM	0.51	-1.5	0.84	0.75	C	Before	0.5	62%	66%
	DC	0.87	-3.5	0.71	0.75	C	Before	0.5	62%	70%
	GM *	0.00	N/a	N/a	0.75	N/a	Absent	0.0	62%	50%
BAND B	MS	0.79	+2.75	0.77	0.80	P	Before	0.5	70%	53%
	DS	0.71	0.0	0.87	0.80	P	Before	0.58	70%	80%
	SA	0.52	-0.75	0.94	0.80	P	Before	0.96	70%	67%
	GT	1.07	-1.25	0.60	0.80	P	Before	1.15	70%	71%
	GR	0.8	-2.0	0.83	0.80	P	Before	0.82	70%	80%
BAND C	DT	0.0	-3.0	0.75	0.82	P	Before	0.58	72%	71%
	LH	0.33	+0.5	0.96	0.82	P	Before	0.58	72%	72%
	GT	0.55	-1.25	0.77	0.82	P	Before	0.96	72%	65%
	PA	0.39	-1.5	0.83	0.82	P	Before	0.5	72%	77%
	GR	0.58	-2.0	0.77	0.82	P	Before	0.58	72%	74%
BAND D	HW *	0.56	N/a	N/a	0.82	P	Before	0.82	70%	72%
	EJ	1.10	-3.7	0.67	0.82	P	Before	0.5	70%	70%
	LM	0.89	-1.2	0.90	0.82	P	Before	1.0	70%	76%
	NL	0.9	-0.6	0.91	0.82	P	Before	0.5	70%	75%
	LA	0.62	N/a	N/a	0.82	P	Before	0.58	70%	59%
	SH	0.64	-1.6	0.83	0.82	P	Before	0.5	70%	68%
	JS	0.77	-1.9	0.78	0.82	P	Before	0.5	70%	Left

Appendix 8i: Example of selective coding: analysis summary matrix 2003-2004

Reference	Coding statements
	Mean of combined Standard Deviations of intra peer assessment (excluding self-assessment) data analysis
03cpqa01	Individual students who exhibit a small Standard Deviation Mean and also tend to receive a high final individual assessment
03cpqa02	Individual students who exhibit a large Standard Deviation Mean and also tend to receive a low final individual assessment;
03cpqa03	Individual students who exhibit a large Standard Deviation Mean also tend to exhibit a large Group Attributes Standard Deviation
03cpqa04	Individual students who exhibit a small Standard Deviation Mean also tend to exhibit a small Group Attributes Standard Deviation
03cpqa05	Slight correlation between individual students exhibiting a small Standard Deviation Mean and their band's Performance assessment mark
04vaqc01	Individual students who exhibit a small Standard Deviation Mean and also tend to exhibit a high Inter Peer assessment
04vaqc02	Individual students who exhibit a large Standard Deviation Mean also tend to exhibit a large Group Attributes Standard Deviation
04vaqc03	No relationship between individual students exhibiting a small Standard Deviation Mean and small Group Attributes Standard Deviations
04vaqc04	Individual students who exhibit a small Standard Deviation Mean also tend to receive a high final individual assessment
04enqc01	Individual students who exhibit a small Standard Deviation Mean have also collaborated in intra peer assessments
04enqc02	Individual students who exhibit a small Standard Deviation Mean have also a small Self Assessment Variance
04enqc03	Individual students who exhibit a large Standard Deviation Mean also tend to have a large Group Attributes Standard Deviation
	Self Assessment variance data analysis
04vaqc05	Individual students who over-estimate their own intra-peer marks also tend to receive a low final assessment mark
04vaqc07	Individual students who under-estimate their own intra-peer marks also tend to receive a high final assessment mark
04enqc04	Individual students who over-estimate their own intra-peer marks also tend to receive a low final assessment mark
	Individual Normalised Difference Coefficient
04vaqc08	Individual students who exhibit a lower Coefficient also tend to receive a low final assessment mark and a large Self Assessment variance.
04enqc06	Individual students who over-estimate/under-estimate their own intra-peer marks have a low Individual Normalised Difference Coefficient
	Band Normalised Difference Coefficient
04vaqc09	Students exhibiting high band averages also tend to receive high final assessment marks
04vaqc10	Students exhibiting low band averages also tend to receive low final assessment marks
04enqc07	Students exhibiting high band averages have also collaborated in intra peer assessments
	Private or collaborative assessment
04vaqc11	Collaborative assessments produced equal distribution of band members' intra-peer marks
04enqc08	Collaboration leads to an equal distribution of intra peer marks among individual band members (Band A)
04enqc09	Collaboration leads to an equal distribution of intra peer marks among individual band members except for a free-rider (Band D)
	Standard deviations of group attributes
04enqc10	Individual students exhibiting a large Group Attributes Standard Deviation have a large Individual Attributes Standard Deviation Mean
04cpqa06	Individual students who exhibit a large Group Attributes Standard Deviation also tend to exhibit a large Standard Deviation Mean
04cpqa07	Individual students who exhibit a small Group Attributes Standard Deviation also tend to exhibit a small Standard Deviation Mean

Reference	Coding statements (continued)
	Standard deviations of group attributes (continued)
03cpqa08	Inconclusive relationship of Standard deviations of group attributes with Inter Peer Band Performance assessments
03cpqa09	Slight correlation between small Standard Deviation of Group Attribntes also tend to receive higher Final assessment mark
03cpqa10	Slight correlation between large Standard Deviation of Group Attributes also tend to receive lower Final assessment mark
04vaqc12	Collaboration in the Intra Peer assessment (Band A), leads to each student sharing the same standard deviation
04vaqc13	Students exhibiting small Standard deviations of group attributes also tend to exhibit small Standard Deviations of individual attributes
04vaqc14	Students exhibiting large Standard deviations of group attributes also tend to exhibit large Standard Deviations of individual attributes
	Assessment conducted before or after receiving performance mark
04vaqc15	All intra-peer assessments conducted before students received their band performance mark
04enqc11	All intra-peer assessments conducted before students received their band performance mark
	Inter Peer Performance Assessment band mark awarded by peer panel
03cpqa11	Inconclusive relationship between Interpeer Assessment, Standard Deviation of Group Attrib & Standard Deviation of Individual Attrib
04vaqc16	Bands exhibiting Inter Peer marks also tend to receive high Band Normalised Difference Coefficients (see previous)
04enqc12	The band who received the highest Inter Peer marks also exhibited a small Individual Attributes Standard Deviation Mean
	Final Individual Grade
04enqc13	Individual students who receive a high final individual grade also tend to have a small Individual Attributes Standard Deviation Mean
	Free Riders
03cpqa12	GM exhibits large Standard Deviation of Individual Attributes and also large Standard Deviation of Group Attributes (Band 2)
03cpqa13	KL exhibits large Standard Deviation of Individual Attributes and also large Standard Deviation of Group Attributes (Band 4)
03cpqa14	LA exhibits large Standard Deviation of Individual Attributes and also large Standard Deviation of Group Attributes (Band 5)
04vaqc17	SK over-estimated in self-assessment (Band A)
04vaqc18	MS over-estimated in self-assessment (Band B)
04vaqc19	GT under-estimated in self-assessment (Band C)
04vaqc20	LA absent for self-assessment (Band D)
04enqc14	LA had smallest Individual Attributes Standard Deviation Mean (Band D)
04enqc15	MS had largest Individual Attributes Standard Deviation Mean (Band B)
04enqc16	GT had largest Individual Attributes Standard Deviation Mean (Band C)
	Shooting Stars
03cpqa15	DC exhibits above average Standard Deviation of Individual Attributes & below average Standard Deviation of Group Attributes (Band 1)
03cpqa16	MS exhibits average Standard Deviation of Individual Attributes and small Standard Deviation of Group Attributes (Band 2)

Appendix 8j: Example of selective coding: integration of the analysis summary matrix 2003-2004



Secondary (grounded in statistically treated primary data)

Mean Correlation of intra-peer and self-assessments by student as measured by normalised coefficients

High (0.90 or higher)	8
Moderate 0.75 or higher)	13
Low (0.75)	6

Consistency of marking agreement for group attributes as measured by standard deviation

Group attributes exhibit a small standard deviation (0.5 or smaller) overall	70
Group attributes exhibit an average standard deviation (0.6 – 1.2) overall	29
Group attributes exhibit a large standard deviation (1.3 or larger) overall	4

Correlation of intra-peer and self-assessments of individual attributes (personal weaknesses) as measured by normalised coefficients

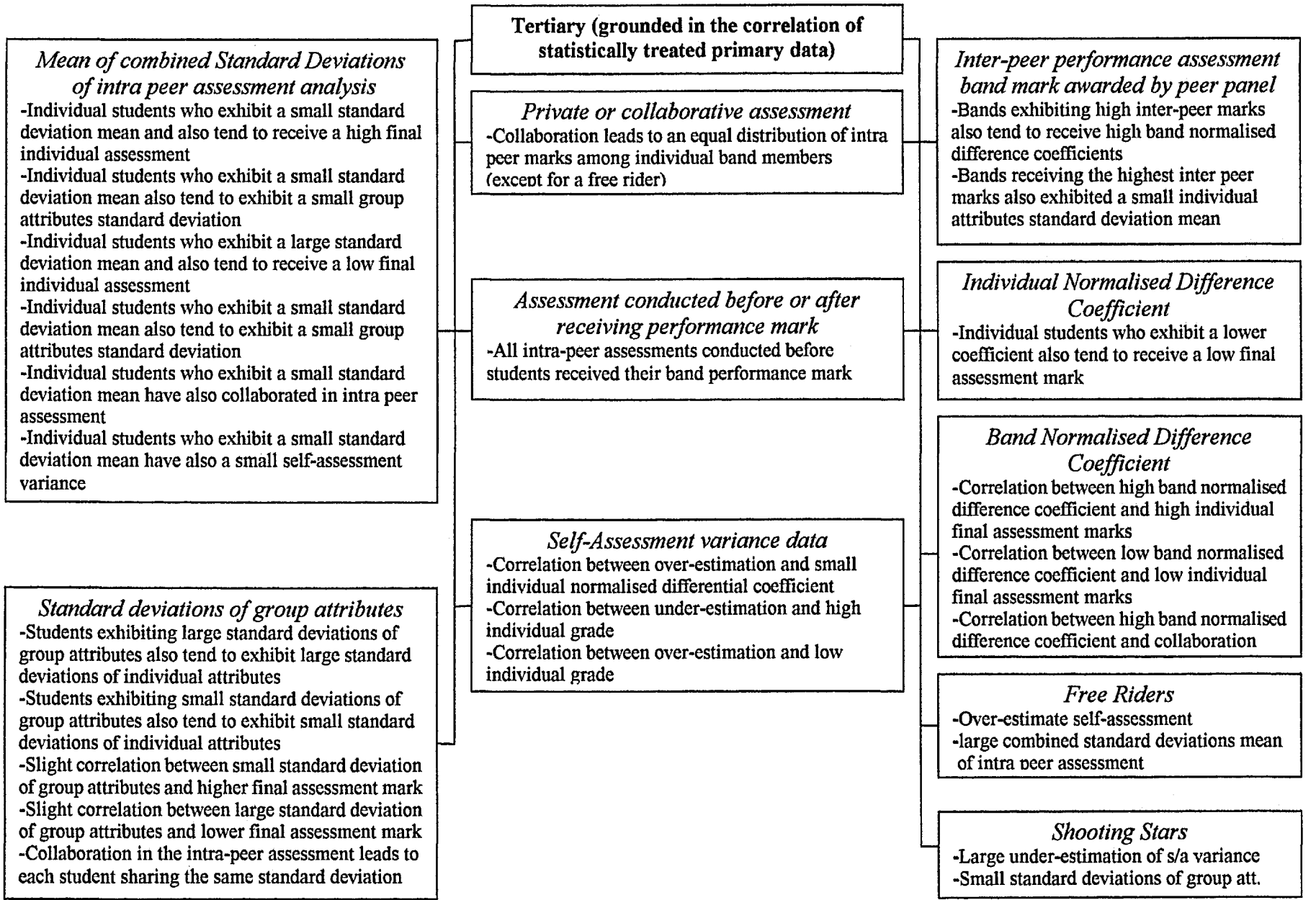
High (0.90 or higher)	32
Moderate (0.75 or higher)	25
Low (less than 0.75)	24

Consistency of marking agreement for personal weaknesses as measured by standard deviation

3 personal weaknesses exhibit a small standard deviation (0.5 or smaller) overall	111
3 personal weaknesses exhibit an average standard deviation (0.6 – 1.2) overall	75
3 personal weaknesses exhibit a large standard deviation (1.3 or larger) overall	18

Comparing consistency of marking agreement between group attributes and personal weaknesses as measured by standard deviation

Number of personal weaknesses having a smaller standard deviation than group attributes	18
Number of personal weaknesses and group attributes having an equal standard deviation	5
Number of personal weaknesses having a larger standard deviation than group attributes	29


Tertiary (grounded in the correlation of statistically treated primary data)
Mean of combined Standard Deviations of intra peer assessment analysis
-Individual students who exhibit a small standard deviation mean and also tend to receive a high final individual assessment
-Individual students who exhibit a small standard deviation mean also tend to exhibit a small group attributes standard deviation
-Individual students who exhibit a large standard deviation mean and also tend to receive a low final individual assessment
-Individual students who exhibit a small standard deviation mean also tend to exhibit a small group attributes standard deviation
-Individual students who exhibit a small standard deviation mean have also collaborated in intra peer assessment
-Individual students who exhibit a small standard deviation mean have also a small self-assessment variance
Private or collaborative assessment
-Collaboration leads to an equal distribution of intra peer marks among individual band members (except for a free rider)
Assessment conducted before or after receiving performance mark
-All intra-peer assessments conducted before students received their band performance mark
Self-Assessment variance data
-Correlation between over-estimation and small individual normalised differential coefficient
-Correlation between under-estimation and high individual grade
-Correlation between over-estimation and low individual grade
Standard deviations of group attributes
-Students exhibiting large standard deviations of group attributes also tend to exhibit large standard deviations of individual attributes
-Students exhibiting small standard deviations of group attributes also tend to exhibit small standard deviations of individual attributes
-Slight correlation between small standard deviation of group attributes and higher final assessment mark
-Slight correlation between large standard deviation of group attributes and lower final assessment mark
-Collaboration in the intra-peer assessment leads to each student sharing the same standard deviation
Inter-peer performance assessment band mark awarded by peer panel
-Bands exhibiting high inter-peer marks also tend to receive high band normalised difference coefficients
-Bands receiving the highest inter peer marks also exhibited a small individual attributes standard deviation mean
Individual Normalised Difference Coefficient
-Individual students who exhibit a lower coefficient also tend to receive a low final assessment mark
Band Normalised Difference Coefficient
-Correlation between high band normalised difference coefficient and high individual final assessment marks
-Correlation between low band normalised difference coefficient and low individual final assessment marks
-Correlation between high band normalised difference coefficient and collaboration
Free Riders
-Over-estimate self-assessment
-large combined standard deviations mean of intra peer assessment
Shooting Stars
-Large under-estimation of s/a variance
-Small standard deviations of group att.

Appendix 9: Example of an agreement form for band determined personal weaknesses

Members of your previous band:

Name of Student _____

Please list three qualities/attributes or 'personal weaknesses' that you as a band agree that _____ should try to develop/improve in their rehearsing.

1.....

2.....

3.....

What is the opposite/negative of each of these three?

1.....

2.....

3.....

I agree to these 'personal weaknesses' to be used as assessment criteria for rehearsing with my next band.

Signed..... Date.....