RADAR
Research Archive and Digital Asset Repository

OXFORD BROOKES UNIVERSITY

Towards a pragmatic and evidence-based approach to the assessment of group work in undergraduate courses: an action research study

Trudy Dee McAuley (2009)

https://radar.brookes.ac.uk/radar/items/05254c04-530b-4953-8d98-5b4c736af853/1/

Note if anything has been removed from thesis: fig. 1.1 on p.9

Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this work, the full bibliographic details must be given as follows:

McAuley, T D (2009) *Towards a pragmatic and evidence-based approach to the assessment of group work in undergraduate courses: an action research study* PhD, Oxford Brookes University

WWW.BROOKES.AC.UK/GO/RADAR

Towards a pragmatic and evidence-based approach to the assessment of group work in undergraduate courses: an action research study

Trudy Dee McAuley

Awarded by Oxford Brookes University

This thesis is submitted in partial fulfilment of the requirements of the award of Doctor of Education

April 2009

BRITISH LIBRARY

IMAGING SERVICES NORTH

Boston Spa, Wetherby
West Yorkshire, LS23 7BQ
www.bl.uk

UNABLE TO COPY AT THE REQUEST OF THE UNIVERSITY

FIGURE 1.1 PAGE 9

Acknowledgements

I would like to thank my supervisors, Elaine Cox, Kathryn Ecclestone and Tatiana Bachkirova for their advice and feedback over the last four years. Also, many thanks to my husband, Andrew, for his support, encouragement and willingness to listen patiently at any time of the day or night.

Abstract

This action research study focuses on *students' perceptions* of group work and its assessment. The research records a four-year study involving 729 first, second and third year undergraduate students. Having started with an exploration into students' perceptions of group work, it identified social loafing as students' key concern, explored ways to alleviate the problem, presented students with a range of options and ended with an examination of students' perceptions of the formative and summative peer assessment process they chose to implement.

Research was undertaken using a mixed model design referred to as "intramethod mixing" by Johnson and Turner (2003, p. 298). This involved the concurrent and sequential use of a single method (questionnaires) that included both qualitative and quantitative components.

As a result of the research a pragmatic and evidence-based approach to the assessment of group work on undergraduate courses is suggested. Also six overarching themes emerged:

- Group work as a social activity that can foster the development of a community of practice.
- Students' lack of knowledge of how to work in groups and need for relevant training.
- Students' dislike of receiving one group mark for group work.
- Tension between fairness, transparency and validity when peers summatively assess peers' contribution to group work.
- The role of formative peer assessment (which should remain anonymous) in assessment *as* learning.
- Tension between the delegation of control over the assessment process and the tutor's role as assessor.

List of Tables and Figures

Tables

Figures

Appendices

Contents

Chapter 9: Students' perceptions of the summative and formative peer assessment process

Chapter 10: Conclusion

Chapter 1: Introduction

Traditionally HE, like other sectors of the education system, has been driven by an imperative to assess individual achievement in a norm-referenced system designed to select the best for the top grades (Grant, 1994, p.127; Ecclestone, 2008). This individualistic ethos has discouraged students from working collaboratively on assessed tasks. However, over the last twenty-five years an interest in summatively assessed group work has increased in higher education. This interest has been driven by various factors.

Firstly, there was a growing realisation, fuelled by criticisms from employers, that graduates needed to be better equipped to work in groups or teams. Cameron et al (2004, p. 2) for example, state: "Much of today's work is organized through teams and requires team collaboration and team working for it to succeed." Research undertaken by TMP Worldwide Research (1998) together with Green's (1990) analysis of advertisements for graduate level posts found that the skill second most frequently requested by employers was the ability to work effectively as part of a team. Teamwork was seen to be vital in a world where an increasing number of jobs "require more effort, knowledge, skills, ways of thinking, ideas and flexibility than one person alone can bring to them" (Levin, 2005, p.11).

Employers also wanted graduates who could collaborate with a range of people with specialisms other than their own. It was claimed that group work could help students hone the social skills they needed to develop successful relationships with others and make this possible. Group work was also seen as a way of developing individual responsibility and task-related transferable skills such as presentation, problem solving, leadership, delegation and organisation, which are necessary for employment (see for example Butcher et al, 1995; Maguire & Edmondson, 2001; and Gibbs, 1994).

Secondly, there was pressure on staff resources caused by an increase in student numbers in higher education. This encouraged academics to explore ways in which students could work in groups without supervision. Academics were also becoming increasingly aware of the role group work could play in improving the quality and enjoyment of the learning experience (Thorley & Gregory, 1994). In addition, there was an attempt to incorporate a range of less threatening teaching and assessment methods, such as group work, into higher education in order to motivate non-traditional students and encourage wider participation (Ecclestone, 2008).

Thirdly, in the UK, national movements such as Higher Education for Capability and the (then) Department of Employment's Enterprise in Higher Education Initiative stimulated much innovation. Both encouraged academic staff to develop or enhance approaches to teaching and learning that were specifically directed at facilitating students' acquisition of personal transferable skills, particularly those, such as the ability to work in a group, relevant to employment.

Despite the increase in the use of group work in higher education, much of the literature on the subject is presented from a tutor's, employer's or policy maker's perspective. There is a relatively small body of research on *students' perceptions* of group work. Indeed, a large number of the studies reported involve discussions of the reliability and validity of tutor versus peer awarded marks for group work (for example, Orpen, 1982; Freeman, 1995; Pond et al, 1995; Falchikov & Goldfinch, 2000) and focus on the *product* of group work rather than the *process* from the students' point of view. This study aims to address the gap in current research by focussing on *students' perceptions* of the *process* of group work and its assessment.

This introductory chapter contains four sections. Section 1 lays out the conceptual framework that underpins my interest in group work as a teaching and learning tool. It also provides background information and

outlines the context in which the study took place. Section 2 defines 'group work' for the purposes of the study. Section 3 outlines the aims of the study together with the research questions involved. It also delineates the action research process by providing an overview in diagrammatic format (Figure 1.1). Section 4 sets out the structure of the thesis.

Section 1: Conceptual framework and context

I am a senior lecturer in education and social psychology in a School of Education in a large university in the south of England. Undergraduates on my modules often work in groups to discuss readings, present different sides of an argument, consider the research undertaken by other group members on a subject and complete tasks in class. They also produce posters, complete projects and make presentations in groups on which they are assessed.

Argyris and Schön (1974) refer to espoused theories that reflect practitioners' genuine beliefs about what underpins their practice. The espoused theory that underpins my practice is social constructivism. I believe that knowledge is not transmitted but constructed by those involved in the learning process (see for example, Glasersfeld, 1989; Winne & Butler, 1994; Biggs, 2003; Gergen, 1995) and that it involves a dialectical process of meaning-making whereby students actively produce rather than passively reproduce meaning (Elliott, 1991). Since I believe this to be the case, my primary concern as a tutor is to create a learning environment that optimises students' opportunities for meaning-making. Vygotsky (1962, 1978) believed that interactions hold the key to meaning-making. Discussion encourages students to articulate their ideas, exposes them to other points of view and encourages the understanding of new perspectives and the creation of meaning (Cheng & Warren, 2000; Webb, 1989 and 1995). My interest in group work developed because it provides an effective context for such meaning-making by increasing the opportunity for interaction between students.

Since group work has become central to my teaching and learning philosophy, I wanted the opportunity to step back and critically evaluate its role and effectiveness. Although evidence collected on module evaluation forms suggests that students appreciate the benefits of group work, I wanted to study their perceptions of group work in more depth to establish what they liked and disliked about it with the aim of improving one or more elements of my own practice and contributing to the knowledge base around group work.

The university in which the research was undertaken operates a modular system, referred to as the UMP (Undergraduate Modular Programme) for undergraduates who are not following professional programmes such as teaching or nursing. In their first year, students on the UMP must take a certain number of modules that are compulsory to their degree programme – usually four. They are then free to select another four modules from over 2,000 'basic' (Year 1, level 4) modules as 'top up' modules to complete their first year programme. Human Communication is a compulsory module (for which I am module leader) for all single honours students on a Communication Studies degree but it is also popular with students throughout the university as a top up module. Since the module recruits many students (usually between 120 and 160) from a wide range of academic programmes, I decided to focus initially on this module in order to elicit the opinions of a large number of students.

Section 2: Definitions

Before beginning a study to explore students' perceptions of group work, it is necessary to define what a group is. The issue is controversial because although many social scientists believe they have defined the concept *group* clearly, their definitions reflect a high level of disagreement concerning the fundamental aspects of what constitutes a group.

According to Johnson and Johnson (2003, p.15) groups can be defined in terms of any one of the following descriptors:

1. Goals – people join together to achieve something they could not achieve on their own.
2. Interdependence – a group is only a group if an event that affects one of the members affects all members.
3. Interaction – a group does not exist unless interaction occurs.
4. Perceptions of membership – people are only members of a group if they perceive themselves to be.
5. Structured relationships – a group only exists if individuals' interactions are structured in terms of roles and norms.
6. Mutual influence – a group exists only if individuals affect and are being affected by each other member.
7. Motivation – a group only exists if individuals are in it to obtain some sort of reward or the satisfaction of personal needs through joint association.

For the purposes of this study, groups are defined in terms of a synthesis of the above descriptors: groups consist of between three and five members who recognise that they are a group, who are working towards a summatively assessed piece of work, are interdependent, interact on a face-to-face basis at least once a week and have some sort of structured relationship.

Ravenscroft et al (1999, p.163) comment that there may be as many definitions of *working* in a group as there are researchers writing about it. The most commonly used terms are "collaborative", "cooperative" and "group-based" learning. Although these terms are often used interchangeably, some authors feel that there is a meaningful difference between cooperation and collaboration. Underwood (2003), for example, focuses on what students *do* in a group. When students are working on their own on sub-tasks which are then pooled, she defines this as "cooperative learning" (p.320); whereas when students work together on

all parts of the task, she defines this as "collaborative learning" (ibid). McWhaw et al (2003) and Ravenscroft et al (1999) on the other hand focus on the curriculum when differentiating between cooperative and collaborative learning. According to them, cooperative learning involves content determined by tutors, whereas collaborative learning is a more democratic process which includes students defining and designing the curriculum (see also Bruffee, 1993).

Ravenscroft et al (1999) and Underwood (2003) also refer to "group-based learning" which they claim is an all-embracing term that covers both cooperative and collaborative learning. For the purposes of this study the more general term "group work" is adopted. This refers to the learning context in which small groups of students (3 – 5) work together to produce a product (for example, a project, poster, or presentation), which is to be summatively assessed. Although students are not normally involved in designing the curriculum, they do have a certain amount of freedom in terms of the nature by which they achieve pre-determined learning outcomes.

Section 3: Aims and research questions

The aim of this study is to explore undergraduate students' perceptions of group work in order to establish what they like and dislike about it, to try to rectify one or more problems they report and ultimately improve practice.

I designed an action research study involving three cycles. The first cycle of research aimed at addressing three initial research questions:

- What are students' perceptions of group work in higher education?
- What do they like about it?
- What problems do they report?

This first cycle of research unearthed key problems relating to what students described as the 'unfairness' of allocating one group mark if one or more group members (termed 'social loafers') did not contribute and led to the following questions:

- Why does social loafing occur?
- If the main cause is lack of identifiability, how can the latter be incorporated in the assessment of group work?
- What strategies of identifying individual contribution to group work exist?

This led to the second cycle of research which involved students choosing peer assessment of contribution to group effort as a way of addressing the problem of social loafing and raised a number of further research questions such as:

- What method of peer assessment would students in my study choose and why?
- What factors need to be taken into account when implementing summative peer assessment?
- What does the literature advise on criterion-referenced versus holistic assessment and which would the students in my study choose and why?
- If the former, who should choose the criteria?
- Should self-assessment be included or excluded?
- Should the allocation of summative marks be kept confidential or not?
- How could summative marks be calculated?
- How could a formative element be incorporated into the assessment process so students benefit from peer feedback?

The final cycle of research involved implementing a peer assessment mechanism whereby students summatively and formatively assessed

each others' contribution to the group product (for example, poster, project or presentation) and led to the final research questions:

- What are students' perceptions of a formative and summative peer assessment mechanism introduced to identify individual students' contributions to a group assignment?
- Does such a mechanism alleviate social loafing?

As a result, the initial research question changed from 'what are students' perceptions of group work?' in cycle 1 to 'what can be done to address the problem of social loafing in group work?' in cycle 2 to 'what are students' perceptions of the assessment of group work when a formative and summative peer assessment mechanism is incorporated?' and 'does such a mechanism alleviate social loafing?' in cycle 3. Addressing these questions resulted in a pragmatic and evidence-based approach to the assessment of group work in higher education.

Due to the "emergent" nature of the research (Lincoln & Guba, 1985, p. 208), this report does not follow the traditional layout of a thesis. Although a short literature review is located in Chapter 2, as is traditionally the case, an additional series of literature reviews were undertaken as each cycle of the action research process uncovered new questions to be addressed. Consequently, subsequent literature reviews are woven into each cycle rather than isolated from the emergent questions they sought to illuminate. This format is designed to make it easier for the reader to follow the complex research journey.

The research process is represented *retrospectively* in Figure 1.1 using McNiff's (2002) model of action research to portray the questions that emerged throughout the research. A more detailed table is available in Appendix 1.1.

Figure 1.1 Simplified version of research questions addressed in the study using McNiff's (2002) action research model

Section 4: Structure of thesis

Chapter 2 explores the advantages and disadvantages of group work identified in the literature. It concludes with an outline of the key theories underpinning the use of group work as a teaching and learning tool.

Chapter 3 explores the epistemological and ontological assumptions underpinning the methodological approach taken. It introduces the main reasons for adopting an action research methodology and explores and critiques different action research models. It also describes the methods

to be used and discusses ethical issues, data collection, data analysis, reliability, validity and generalisability.

Chapter 4 addresses the first research questions in the preliminary cycle of action research: What are students' perceptions of group work; what do they like about it and what problems do they experience? It outlines the method used before discussing the findings in relation to the literature. It concludes by outlining the problems identified together with action to orientate the second cycle of research.

In Chapter 4 the issue of social loafing is identified as a key problem in group work by students on a first year module. Chapter 5 explores the theoretical explanations for social loafing and establishes the importance of ensuring individual accountability in order to deter social loafing. The chapter offers a brief overview of the methods of assessing individuals' contribution to group work before considering the three main methods in detail and discussing students' responses in the literature.

Chapter 6 discusses the findings from research with two samples of students in terms of their choices of methods of assessing group work. It then focuses on the choices made by social loafers and discusses possible reasons for their choices. It suggests questions that need to be resolved before the third cycle of the action research process can be undertaken.

Research with two groups of students reported in Chapter 6 led to the decision to implement summative peer assessment in order to identify individual contributions to group work. Chapter 7 contains a literature review, which seeks answers to a number of questions such as whether summative peer assessment should be based on one holistic mark or on a number of pre-specified criteria; who should decide on the criteria to be used if criterion-referenced assessment were chosen; whether self-assessment should be included or excluded; and whether or not the allocation of summative marks should be confidential. It also explores

how a formative element could be incorporated so that students receive the benefit of feedback to facilitate improvement of their group working skills and discusses whether or not this feedback should be anonymous.

In Chapter 8 students are presented with the opportunity to make their own choice in terms of how peer assessment should be implemented. The chapter also describes the evolving process for deciding on the criteria for peer and self-assessment of contribution to the group during modules and explains how assessment was implemented.

Chapter 9 explores students' feelings about being involved in the process of setting criteria and reports their responses to the formative and summative peer assessment process described in the last chapter. It then discusses the findings in relation to the literature. The chapter concludes with an examination of the extent to which peer assessment of contribution alleviates the problem of social loafing.

Chapter 10 draws together the main findings in terms of themes. It includes a reflection on the process together with the nature of assessment and the dilemmas it produced. The chapter concludes with a discussion of what I consider to be the strengths and limitations of the research and suggests questions to orientate the next cycle of my continuing action research.

Chapter 2: Literature Review

The purpose of this initial literature review is to explore the advantages and disadvantages of group work both as a general pedagogy and a summative assessment tool and to make explicit the theories that underpin its use. Section 1 discusses the advantages of group work identified in the literature. Section 2 explores the key theories underpinning group work. Section 3 explores the disadvantages of group work. Section 4 contains the conclusion.

Section 1: Advantages of group work

Supporters of group work argue that it helps develop skills for life including the world of work and that it can be an effective teaching and learning tool.

Development of skills for life

In Chapter 1 it was established that employers expect graduates to be able to work effectively as part of a group. According to a number of sources (for example, Falchikov, 1993; Maguire & Edmondson, 2001; Harvey et al, 1997; Jaques, 1984; Cheng & Warren, 2000; Underwood, 2003; Brown & Knight, 1994; Goldfinch and Raeside, 1990; Mello, 1993; and Lejk & Wyvill, 1996) group work prepares students for the real world of work where team-working skills are in demand.

Supporters of group work as a pedagogy claim that it helps students hone the interpersonal/social skills they need to develop successful relationships with others (for example: Johnson et al, 1991; Bourner et al, 2001; Maguire & Edmondson, 2001; Oldfield & MacAlpine, 1995; Gibbs, 1994; McWhaw et al, 2003; Mello, 1993; Lejk and Wyvill,1996; Jordan & LeMetais, 1997; Gillies, 2000). It is argued that group work can also help students develop more effective intergroup relations. Slavin (1989b), for example, claims when students of different racial or ethnic background

are involved in group work, they are more likely to learn to like and respect each other.

According to Butcher et al (1995), Maguire & Edmondson (2001) and Gibbs (1994) group work can help to develop individual responsibility and task-related transferable skills such as presentation, problem solving, leadership, delegation and organisation which are necessary for employment.

Effectiveness as a teaching and learning tool

Group work has been widely accepted by university researchers as an effective teaching and learning tool (Conway et al, 1993; Earl, 1986; Jaques, 1984, Mello, 1993). Li (2001) reports that group work is relevant to a wide spectrum of disciplines, for example surgery (Burnett and Cavaye, 1980), mathematics (Earl, 1996), business (Freeman, 1995), biosciences (Garvin, 1995), civil engineering (Rafiq & Fullerton, 1996) and optometry (Conway et al, 1993). Goldfinch and Raeside (1990) argue that group work can cut down on marking while Maguire & Edmondson (2001) suggest that it can help students make friends, especially at the beginning of a course. Burnett & Cavaye (1980) found group work developed feelings of responsibility for other members of the group while Underwood (2003, p.320), Bennett, (1991) and Gillies (2000) claim it can help to improve relationships between peers and Underwood found it improved relationships between tutors and students (see also Reid et al, 1982).

According to Lazarowitz & Karsenty (1990) in Gillies (2000) and Johnson et al (1991) working in groups can also increase self-esteem and facilitate learning in very large classes (Butcher et al., 1995, p.165). It can provide a more interesting learning environment than traditional lectures (Butcher et al., 1995, p.165) and be a more enjoyable social experience (Falchikov, 1993, p.276; Levin, 2005, Cheng & Warren, 2000, p.243).

McDowell (1995, p.307) reports that students in her study felt positively about working within a group in terms of sharing workload, developing and refining ideas, drawing on individual strengths and providing mutual support.

Group work also allows the development of more comprehensive assignments than is possible for individual assessments according to Mello (1993) and Lejk and Wyvill (1996). It maximises available resources (Knight, 2004) by bringing more skills and talent to bear on difficult tasks. McDowell's (1995, p.307) students, for example "saw fellow students as a resource to be tapped". When there are a high number of good quality interactions between students working in a group, the possibility that group members will acquire relevant information to achieve the set task is also increased according to Sharan & Shacher (1994).

Students exhibit intrinsic motivation on group project work particularly if they see the task as realistic or relevant, according to McDowell (1995). She found that although students in her study often complained about the hard work involved "it often seemed to be the case that struggle, effort and facing difficult challenges contributed to a sense of achievement" (McDowell, 1995, p.308).

Maguire & Edmondson (2001) claim that the security of working as part of a group can be a helpful first step in the progression to independent and autonomous learning. The quality of learning can also be improved by peer support (Underwood, 2003 p. 320), pressure (Maguire & Edmondson, 2001) and improving motivation to learn (Sharan & Shaulov, 1990 in Gillies, 2000). Conway and Kember's (1993, p.45) students reported that they found group projects "more effective than lectures for most aspects of learning." Levin (2005) claims group work forces students to confront questions of values, judgement, risk and quality of information. Gillies (2000) found that students who had been trained in group work techniques used higher level cognitive strategies in their interactions and obtained higher marks than their untrained counterparts.

Sharan (1980) postulates that students perform more effectively when working in small groups than being taught as a whole class. It is also claimed that group work can increase the development of critical thinking skills (McWhaw et al, 2003; Searby & Ewers, 1997) and help develop an "issue-oriented" way of thinking (Levin, 2005).

Not only is it claimed that group work allows students to gain firsthand experience of theoretical aspects of group dynamics and group processes (Mello,1993; Lejk and Wyvill, 1996) but it also engages students actively in the learning process (according to Ravenscroft et al, 1999; Underwood, 2003; Mello, 1993; Lejk and Wyvill, 1996) if used effectively (Brown & Knight, 1994; Livingstone & Lynch, 2000). The theoretical premise is that the more actively learners process information through interaction, the more likely they are to learn.

Group work can provide "a vehicle for decision making that permits multiple and conflicting views to be aired and considered" according to Gatfield (1999, p.366). It also helps students learn to think and solve problems "through collective experiences and shared thoughts" according to Underwood (2003, p.320). (See also Tolmie and Boyle, 2000.)

It is argued that group work can facilitate the co-construction of knowledge and that it accords with Vygotsky's belief that interactions produce new understanding for those involved. Discussion in group work exposes students to other points of view, encourages justification of ideas, resolution of disagreements and understanding of new perspectives (Cheng & Warren, 2000; Webb, 1989; Webb, 1995, p.244). It also facilitates giving and helping interactions among students (Cowie & Rudduck, 1990 in Gillies 2000 p.98; Webb, 1989; Webb et al 1995). Research has shown that it is the interactions that occur in groups that facilitate learning (Bennett, 1991, Webb, 1989,1991,1995).

> A primary motivation for putting students into groups is the opportunity for students to help each other learn. Students can learn from other students by giving and receiving help; by

> recognising contradictions between their own and other students' perspectives, seeking new knowledge to resolve those contradictions, and constructing new understandings from them; and by internalising problem-solving processes and strategies that other students use or that are created jointly with others (Webb et al, 1995, p.406).

Bales (1970) found that when students interact with each other and seek information, they generally attract five times as many information exchanges for each request, hence increasing the possibility that group members will exchange more relevant information. It is this information that is believed to be related to learning gains. Sharan & Shachar's (1988) research with children showed that increased participation in group discussions resulted in more intellectually valuable contributions to these discussions and the quantity of speech included high level contributions and not just a larger flow of words. Webb (1989, 1991, 1992) found that providing explanations in response to requests for help benefits both the *receiver* and the *giver.* Webb et al (1995, p.406) suggest that students are often more aware than their teachers of what other students do not understand, can focus on the main aspect of the problem and provide explanations that can be more easily understood. Webb (1989, p.35) explains what sort of explanations *recipients* might benefit from:

> ..... relating new material to what students already know, using multiple representations to explain a concept .... providing detailed justification of each step in the problem solving process, using specific examples and translating unfamiliar vocabulary into familiar terms.

Giving explanations also benefits the *giver*; it requires the reorganisation and clarification of the materials to be learned which may help the explainer to develop new perspectives, construct more elaborate cognitive understandings, and, in so doing, often learn the material better than before (Doise & Mugny, 1984; Wittrock, 1990). Unsolicited explanations can also be of benefit. According to Gillies & Ashman (1996) students working in groups often develop an intuitive sense of others'

needs and offer the necessary help. Forman (1989) refers to help that is timely, relevant to the student's needs given by a more able student as "proleptic instruction" (p.57). Shachar & Sharan (1994) found that when students were involved in group activities that encouraged them to use language as a medium with which to represent their ideas, they used more sophisticated and a more divergent range of thinking strategies. These strategies ranged from thinking about the abstract and hypothetical through to the more extensive strategies required for organising and structuring their thoughts and the ideas of their peers with whom they are communicating. Gillies and Ashman (1996, pp.99 – 100) point out that these studies suggest, in effect, that high-level group tasks "promote greater levels of interaction and higher level thinking behaviours."

Perhaps the most important claim is that group work promotes student achievement (Johnson et al, 1991; Johnson, 1992; Bourner et al, 2001; McWhaw et al, 2003 who refer to Cockrell et al, 2000; Hiltz, 1998; Webb et al, 1995; Johnson et al, 2000; Slavin, 1992). Webb et al, for example, found that the level of "constructive activity" was a strong predictor of achievement. Constructive activity refers to solving or explaining how to solve problems using concepts stated or implied in the explanations received from group members (1995, p.406).

Research undertaken by Knight (2004) compared 70 second-year undergraduate students' perception of their performance with their actual performance in individual and group assessment. He found that although students tended to think they performed better in individual assessment, they actually performed better and achieved greater perceived development of key skills as a result of assessed group work.

According to Slavin (1983) cooperative learning techniques produce noticeable gains in student learning in 80% of studies. Brown (2000) claims that organising students into cooperative learning groups produces consistently better student achievement. These studies add weight to Johnson et al's (1981) findings from a meta-analysis of 122 studies which

found that co-operation promoted higher achievement and increased productivity across all subject areas.

One method of assessing group work is to involve the students in the assessment process. This results in the opportunity to strengthen skills in peer and self-assessment. Brown & Knight (1994, p.61) claim that it is important that students experience peer assessment so they are exposed to situations in which they are required to respond sensitively and perceptively to peers' work.

Section 2: Theories underpinning the argument in favour of group work

The positive effects of group work can be explained by motivational theories and social cohesion theory (Slavin, 1989a; McWhaw et al, 2003; Ravenscroft et al, 1999).

Social cohesion theory posits that if group members develop positive relationships with each other they will be concerned for each other's welfare and will work hard to do well (Ravenscroft et al, 1999).

Motivational explanations concentrate on explaining why students are interested in learning in groups (for example: Slavin, 1992; Abrami et al, 1995; Ames, 1984, Cohen, 1994, Sharan & Sharan 1992) and are based on the work of Kurt Lewin and Morton Deutsch on social interdependence which has been extended by the Johnson brothers more recently. Johnson & Johnson (1994) suggest that social interdependence theory motivates students to engage in promotive interactions in order to achieve mutual goals. The theory posits that motivation determines action; students are motivated by interdependence in group work; they realise that they sink or swim together; it is to everyone's advantage that the group does well. However, proponents of this theory are not always in agreement over the role of individual and group incentives. Some believe

that students can be intrinsically motivated by the task alone; others believe that extrinsic rewards (such as marks/grades) are essential to student motivation (Ravenscroft et al (1999, p.165). In order to ensure that both extrinsic and intrinsic learners engage in group work it is important to develop interesting group tasks that are then assessed.

Apart from motivational theories and social cohesion theory, two additional theoretical frameworks underpin group work: Piaget's (1965) constructivist theory of cognitive conflict and Vygotsky's (1962, 1978) social constructivist theory.

Piaget's theory of cognitive conflict

Piaget argued that when individuals encounter new information which does not fit into their existing mental schema, the contradiction causes "disequilibrium". The learner then has to "accommodate" the new information by modifying his or her current understanding of aspects of the world in order to achieve a new "equilibrium" or alignment of perspectives (Falchikov, 2001, p. 86). Learners, in other words, actively construct their own knowledge. Gillies and Ashman (2003, p.12) explain cognitive conflict as when students are:

> forced to re-examine their understandings and perspectives in the light of contradictions that occur from interacting with others. When this happens, [students] reflect on their own understandings, seek additional information to clarify the contradictions, and attempt to reconcile their perspectives and understandings to resolve any inconsistencies. Cognitive conflict is a catalyst for change as it motivates [students] to reassess their understandings of the world and to construct new ones that fit better with the feedback they are receiving.

Group work provides the opportunity for cognitive conflict when others have different opinions (McWhaw et al, 2003). Students have to wrestle with conflicting information they receive from their peers and incorporate elements of it into their current understanding. Dansereau (1985) and

Webb (1989) refer to this process as cognitive restructuring. Piaget believed that group work was an effective way to encourage a decentred perspective allowing the learner to consider multiple points of view.

Vygotsky's theory of social constructivism

Central to Vygotsky's thinking is that social interaction is key to cognitive development. He believed that knowledge construction and the transformation of different perspectives into personal knowledge resulted from cooperative efforts to understand the world. Peer collaboration can increase development through the zone of proximal development which is defined as:

> the distance between the actual development as determined by individual problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers (Vygotsky, 1978, p.86, in Gillies and Ashman, 2003, p.12).

Although Piaget's theory focussed on cognitive development in terms of the learner as an individual, Doise & Mugny (1984) extended Piaget's research by undertaking a systematic empirical investigation into how social interaction affects individual cognitive development. Doise and Mugny built on the Piagetian idea of centration in the belief that it is through interacting with others and coordinating one's approach to reality with that of others, that learning takes place. According to Reimann & Spada (1996, p.191):

> individual cognitive development is seen as the result of a spiral of causality: a given level of individual development allows participation in certain social interactions which produce new individual states, which, in turn, make possible more sophisticated social interaction and so on.

Doise and Mugny's theory of cognitive elaboration stresses the importance of group work. The claim is that articulating one's ideas to others improves learning. This theory assumes that learners are

intrinsically motivated, i.e. the task itself is assumed to be inherently interesting enough to ensure students' engagement (Slavin, 1989a; Ravenscroft et al, 1999). It has led to Ryan's (1997) principle of multiplicity which states that since it is unlikely that two people will see things the same way or have the same solution to a problem, interaction is advantageous for learning.

Working in a group provides the opportunity for socio-cognitive conflict i.e. conflict between different answers based on different centrations. The social dimension can be seen as a catalyst for resolving disequilibrium by transcending the different centrations to arrive at a more advanced decentred solution (Reimann & Spada 1996, p.191).

Piaget's constructivism emphasises that individuals learn best when they actively construct knowledge and understanding. Vygotsky's social constructivist approach emphasises the social contexts of learning and that knowledge is mutually built and constructed. Vygotsky's social constructivist model is a social learner embedded in a sociohistorical backdrop. Moving from Piaget to Vygotsky, the conceptual shift is from the individual to collaboration, social interaction and sociocultural activity. Piaget believed that students construct knowledge by transforming, organising and reorganising previous knowledge and information. Vygotsky believed that students construct knowledge through social interactions with others. The content of this knowledge is influenced by the culture in which the student lives, which includes language, beliefs and skills (Santrock, 2001).

The implication of Vygotsky's model is that tutors should create many opportunities for students to learn with the tutor and with peers in co-constructing knowledge. In both Piaget and Vygotsky's models, tutors serve as facilitators and guides rather than directors and moulders of learning. As Gillies and Ashman (2003, p.70) point out, since the:

> onus for knowledge construction, information researching and product creation rests increasingly with student groups, the role of

the teacher changes from the "sage on the stage" to the "guide on the side".

Section 3: Disadvantages of group work

Despite the many benefits of group work put forward by its supporters, group work also harbours a number of disadvantages. McDowell (1995, p.311) points out that in group work there is often potential conflict between using individuals' strengths and abilities to achieve a better product and enabling group members to develop skills in areas which they most need to develop. This leads to a tension between learning and producing the assessment "product," with a student who is keen to strengthen areas of weakness jeopardising the chances of getting a good mark for the whole group.

She also points out that formal, summative assessment results in students playing safe and avoiding risks whereas "arguably most personal breakthroughs in learning require an element of risk-taking" (p.312). She suggests however that giving credit to individual progress would require a sea change in higher education which tends to focus on notionally fixed standards of achievement rather than the amount of progress made by individuals within the system.

Phipps et al (2001, p.14) claim that many students are apprehensive about group work when they are first introduced to it and Helms & Haynes (1990, p.9) list the many problems which can lead to dysfunctional groups and, as a result, negative learning experiences for the students involved. These problems include disagreements about task purpose or direction; interpersonal conflict and manipulation of the group by various personality types such as an 'aggressor' or 'blocker'; lack of mutual trust or support and incongruent goals or expectations.

Garland (1994, p.418) is one of many to acknowledge the fact that assessing group work is a "thorny problem" and that it raises three critical

questions. The first question is *what should be assessed* (Garland 1994, CSHE 2002): the product, the process or both? Assessing the process involves two main problems: firstly, it lacks the tangible nature that a product (such as a poster, presentation or project) usually possesses, and secondly, tutors are not normally present when the process of group work takes place. McDowell (1995) explains that this is probably why most tutors feel more confident assessing the product rather than the process. However, as Gibbs (1995) points out, focussing only on the product undermines the important role process plays in group work. An increasing number of lecturers assume that if the product is not assessed at all, students lack the incentive to take the group work seriously in the first place.

Assessment, therefore, requires not just attention to outcomes but also to the process that leads to those outcomes (Krause & Popovich, 1996; Jacques, 2000). As Strachan & Wilcox (1996, p.348) state: in group work the "process must be stressed as equally important to product, for it is in the process that the actual learning takes place."

The second critical question is *who should do the assessing*. Although tutors may be well positioned to assess the product, since they are not present during the process, they are not in a position to assess this element. The third question is then, if students are in the best position to undertake the summative assessment of process through some sort of peer assessment mechanism, *how should this be structured*. The idea of involving students in assessment is rejected by a number of academics.

Ecclestone (2008), for example, questions the legitimacy of students' summative assessment if the stakes of assessment are high. Topping (1998) suggests that students might abuse the power peer assessment gives them. Questions over the reliability and validity of students' marking are raised by others for example McDowell (1995), Topping (1998) and Maguire and Edmondson (2001).

Students in some studies report finding peer assessment difficult (for example, Falchikov, 1995; Mowl and Pain, 1995) or uncomfortable (Osmond & Merry, 1996). Ballantyne et al (2002) and Osmond & Merry (1996) cite students' lack of confidence in both their own and peers' abilities as assessors as a key issue in peer assessment which would justify Ecclestone's scepticism with their involvement. Students in Osmond and Merry's (1996) study reported feeling uncomfortable with peer assessment because they felt unqualified to mark others' work and some felt sceptical about the worth of peers' comments. McDowell's (1995) students expressed concerns about their ability to provide constructive feedback and mark fairly, and although they were prepared to participate in the process, they wanted staff to provide additional feedback.

Scoffield & Brindley (1998) found that more than 50% of students in their study believed that assessment should be the tutor's responsibility (see also Searby and Ewers' 1997 study). In his review of peer assessment studies carried out between 1980 and 1996, Topping (1998) identified problems with some students' reluctance to accept feedback from peers and their unwillingness to take responsibility for the process.

Topping et al's (2000) students rated the cognitive challenge and strain of peer assessment as one of its least liked features while Cheng and Warren (1997) report that although their students agreed in principle with peer assessment, most were not supportive of first-year students being involved. Furthermore, some students found it difficult to award low marks to peers even when they were deserved (in Ballantyne et al, 2002).

Scoffield & Brindley's (1998) students found it difficult to avoid personal bias because they often felt more favourable towards their friends. Students in their study questioned their own worth and therefore felt that it was difficult to take their marking role seriously. They also encountered difficulties in the interpretation of assessment criteria due to their lack of experience of assessment.

Students also complain about the time involved in group work particularly if expected to undertake the process outside class time (McDowell, 1995). The process is also time consuming for tutors (Brindley and Scoffield, 1998; Cheng and Warren, 1999, Searby and Ewers, 1997) in terms of setting up the group work in the first place, monitoring it and devising a peer assessment system.

However, despite the above problems, it can be argued that since the ability to work as part of a team is the second most important soft skill required by employers (Green, 1990; TMP Worldwide Research, 1998) students need to be given the opportunity to learn how to deal with the problems inherent in group work.

Section 4: Conclusion

There is a large body of literature that argues the benefits of group work. Johnson et al (1991) claim that between 1900 and 1991 over 675 studies of group work have been undertaken which show that it results in higher achievement, more effective interpersonal relationships and higher self-esteem than individualistic or competitive learning. There are also a number of disadvantages of group work related to summative assessment and the amount of time it takes.

Livingstone and Lynch (2000) note, however, that there is little research on non-traditional teaching methods such as group work from the perspective of the students involved and Peters (1996, p.48) claims that: "Discussion of the issue has largely remained confined within the walls of the academic office, with those most affected effectively silenced." This study aims to address this omission by focussing on students' perceptions of group work and its assessment in a higher education setting.

Chapter 3: Methodology

In this chapter I explain the research design created in order to explore students' perceptions of group work. The chapter is divided into five sections. Section 1 explores the epistemological and ontological assumptions underpinning the methodological approach taken. Section 2 introduces the main reasons for adopting an action research methodology. Section 3 explores and critiques different action research models and presents the action research model adopted for the purposes of this research. Section 4 describes the methods to be used and discusses data collection, data analysis, reliability, validity, generalisability and ethical issues. The chapter concludes with Section 5, which summarises the main points.

Section 1: Epistemological and ontological assumptions underpinning methodological approach

Burrell & Morgan (1979) Cohen & Manion (2004), Cohen et al (2007) and Miles and Huberman (1994) suggest that the choice of educational research topic and the formulations of research questions are largely determined by which epistemological assumptions underpin a researcher's view of knowledge. This current study is located within a social constructivist epistemology because I believe that knowledge is not given but constructed by those involved in the learning process (Glasersfeld, 1989; Winne & Butler, 1994; Biggs, 2003; Gergen, 1995).

It is traditional to differentiate between the conventional (positivist) and alternative (interpretivist) paradigms when considering methodological approaches (Guba and Lincoln, 1988). Indeed, much of the literature on research methodology presents the idea of two armed camps locked in battle (see for example, Fay, 1996; Morgan, 2007; Willig, 2001).

Within the positivist camp the nature of reality is seen in terms of a realist ontology that assumes the existence of an objective external reality

independent of human perception. The researcher is seen as capable of being value free, an objective outsider who, with care, need not contaminate the research. Such a view requires the researcher to adopt a disinterested approach, to 'self-empty' (Fay, 1996, p.202) to bracket his or her personal perspectives and to disappear to the extent that he/she becomes an objective recording device (and postulates that this is possible). The interpretivist camp reflects a constructivist epistemology and accepts that an objective separateness cannot be maintained between researcher and participants and that findings will reflect the values of the researcher to some extent.

However, Fay (1996, p.241) warns against the 'pernicious dualism' inherent in dichotomies (such as the choice between positivism and interpretivism). Adherence to such dualism leads to what Kuhn referred to as 'incommensurability' (cited in Fay, 1996, p.80; Morgan, 2007, p.58) i.e. the belief that each paradigm has such radically different assumptions about the nature of reality that it is impossible to translate research between these paradigms. Morgan (2007) is one of an increasing number or researchers who do not hold with this schism or 'incompatibility thesis' (see also Miles & Huberman, 1994; Tashakkori & Teddlie 1998, p.4). Instead Fay's 'critical intersubjectivity' (Fay 1996, p.212) reflects my belief that although there is no one correct way to view social reality, if we subject our research to scrutiny, do not hide behind an illusory façade of disinterestedness, are explicit in our descriptions and responsive to the evidence as best we can determine, we can produce a view of social reality (albeit provisional and constructed) which helps our understanding of it.

Fay's critical intersubjectivity as a methodological approach accords with a social constructivist epistemology since both stress the construction of intersubjective meaning. It seems to me to offer an essentially pragmatic approach to educational research. The three core tenets of pragmatism, according to Morgan (2007, p.66) are 'lines of action', 'warranted assertions' and the emphasis on 'workability'. 'Lines of action' derive from

the work of William James and George Herbert Mead and refer to decisions which lead to action; the notion of 'warranted assertions' derives from the work of John Dewey and refers to the need for action to be based on sound reasons which can be justified under scrutiny; and the emphasis on 'workability' derives from the work of William James and John Dewey and refers to the need for proposed action to be achievable and fit for purpose in the real world.

My approach to research is based on the belief that the dichotomy of positivist and interpretivist approaches is simplistic and unnecessary and that one-sided paradigm allegiance is limiting (a view supported by others, see for example Swann & Pratt, 1999; Morgan, 2007; and Johnson & Onwuegbuzie, 2004). Miles and Huberman (1994) argue that each approach adds a meaningful layer without necessarily contradicting the other, in any case. Being pragmatic allows one to eschew methodological orthodoxy in favour of "methodological appropriateness" (Patton, 2002, p.72) and instead concentrates on lines of action, warranted assertions and workability (Morgan, 2007). I planned to engage in action research to ascertain students' perceptions of group work and negotiate a workable, intersubjectively created intervention to address one of the problems identified by students. I planned to warrant my assertions for action by using a large sample of students to ascertain what that action should be.

Section 2: Rationale for adopting an action research approach

This section includes a discussion of the four main reasons why an action research approach was chosen.

1. A social constructivist epistemology resonates with the pragmatic methodological approach outlined by Morgan (2007). Both stress the view of knowledge as a dialectical process of meaning-making. Action research was chosen because it involves a pragmatic approach to problem-solving based on the views and interpretations of the participants

(in this case students) involved in the enquiry (Winter, 1989; Zuber-Skerritt, 1992, p.12- 14; Berg, 2004, p.196; Whitehead & McNiff, 2006).

2. Kolb (1984) claims that the construction of knowledge involves a cyclical learning process. According to constructivists, we learn on the basis of observing and reflecting on concrete experience, forming abstract concepts and generalisations and then testing the implications of these concepts in new situations. This leads to new concrete experience and the beginning of a new cycle (Lewin, 1951; Kolb, 1984; Cowan 2006). Since my aim is to improve my practice, which, in turn, assumes active participation in the learning process on my part, I needed a research process that captured and echoed the cyclical nature of the learning process itself.

There are two key elements of Kolb's action research cycle which were particularly relevant: firstly, the here-and-now and secondly, feedback. Improving practice involves the here-and-now. Students were currently engaged in group work and the aim of the proposed research was to explore their views of it and ultimately improve my own teaching and assessment practice.

Secondly, Kolb took into account Lewin's focus on feedback. Lewin (1951) believed that organisational and individual ineffectiveness could be blamed to some extent on a lack of adequate feedback processes resulting in an imbalance between reflection and action. Either people decided to act rather than gather information or became sidetracked by data collection and analysis and forgot to act. Lewin saw the action research process as a way of avoiding the problem of focussing on one to the detriment of the other (Kolb, 1984, p.22). In my research, feedback was a key element; it was the feedback from students in terms of their response to questions about their experience of group work that dictated future action. It was also important that both action and reflection were highlighted since action without reflection would result in ineffective strategies being implemented and reflection without action would not

benefit students in terms of the implementation of tangible improvements leading to a more positive experience of group work.

An action research approach was used because it involves a spiral of cycles (Lewin, 1952; Kolb, 1984; Elliott, 1991; Robson, 1993; McNiff, 2002; Berg, 2004) and embodies the cyclical nature of the learning process. Kemmis and McTaggart (1988b, p.10) describe the action research cycle as involving the development of a plan of critically informed action to improve a situation; action to implement the plan; observation of the effects of the action; and reflection on these effects as a basis for the next cycle of planning, acting, observing and reflecting. Action research is not just a methodological approach to research but also a learning vehicle in itself, an intrinsic, integral part of professional practice (Winter, 1989, p.4).

The aim of my research was to elicit students' perceptions of group work, identify the major problem(s) they reported, implement strategies to address one of these problems and evaluate students' responses to any interventions. This spiral of cycles is apparent in the cyclical nature of the research design detailed below:

- **Reflecting** on module evaluation forms and questionnaires completed by students. These included open questions asking students to reflect on the benefits and problems they experienced as a result of working as part of a group.
- **Planning** what to do to address one of the problems identified by students.
- **Acting** to research the literature and implement an intervention (suggested by students) to alleviate the problem.
- **Reflecting** on students' responses to the intervention.

3. The third reason for choosing action research was that it is practical and aims to *improve* practice; it is ecologically valid. According to Lewin

"research that produces nothing but books will not suffice" (1948, p.41). Action research is described by McKernan (1996, p.5) as "systematic self-reflective scientific inquiry by practitioners to *improve* practice" and is defined by Carr and Kemmis (1986, p.162) as:

> A form of self-reflective inquiry undertaken by participants in social situations in order to *improve* the rationality and justice of their own practices, their understanding of these practices, and the situations in which these practices are carried out. (My italics.)

Kemmis and McTaggart (1988b, p.10), Denscombe (1998) and Swann & Ecclestone (1999) describe action research as the development of a plan of critically informed action to *improve* a situation. McKernan (1996, p.3) claims the aim of action research is to "solve the immediate and pressing day-to-day problems of practitioners." Carr and Kemmis (1986, p.165) claim action research aims at improvement in three different areas: firstly, improvement of practice; secondly the improvement of the understanding of the practice by the practitioners involved and, thirdly; the improvement of the situation in which the practice takes place.

Although there are different opinions as to the precise role theory plays in action research (discussed below), the *practical* nature of action research is seen as key. Action research is a means of solving practical educational problems by intervening in practice and determining the changes that need to take place (Brock-Utne, 1988, p.253). It involves the integration of some practical outcomes related to the actual lives of participants in the research project (Berg, 2004, p.196).

According to Elliott (1978), action research improves practice by developing the practitioner's capacity for discrimination and judgment: it helps to develop practical wisdom, which is "grounded in reflective experiences of concrete cases" (1991, p.53). As a philosopher, concerned with how theory related to practice, he introduced the notion of practical theory and the testing of these practical theories in the real world:

> Action research is concerned with the everyday *practical* problems experienced by teachers, rather than the 'theoretical problems' defined by pure researchers within a discipline of knowledge (Elliott, 1978 p.121).

According to Burns (2000, p. 450) action researchers are "not immediately concerned with adding more "truth" to [the] body of educational knowledge which appears in articles and books" but are "interested in the improvement of the educational *practices* in which they are engaging – how to do their jobs better." Elliott (1991, p.49) agrees: "the fundamental aim of action research is to improve practice rather than to produce knowledge."

I disagree with Elliott and Burns and suggest, like Kemmis and McTaggart, that action research makes possible a way of working which links theory and practice into the one whole: "ideas-in-action" (1988b, p.6). Educational practice can be the "*starting point* for the underlying theory of educational action research" (Brock-Utne, 1988, p.253) and can be "of theoretical importance to the advancement of knowledge in the field" (Zuber-Skerritt,1992, p.12-14) while also leading to practical improvements during and after the research process.

4. The fourth reason action research was chosen was because of the participative role undertaken by both researcher and participants. The action researcher is not considered to be an outside expert conducting an enquiry with subjects but a practitioner carrying out research with and for the people concerned with the practical problem and its actual improvements (Corey, 1949; Zuber-Skerritt, 1992, p.12- 14; Winter, 1989; McKernan, 1996; Whitehead & McNiff, 2006).

Carr and Kemmis (1986) stress the key role of participation: "Those involved in the practice being considered are to be involved in the action research process in all its phases of planning, acting, observing and reflecting" (p.165). Berg (2004, p.196) underlines the importance of "the active engagement of individuals traditionally known as subjects as

participants and contributors in the research enterprise" and Brock-Utne states:

> It is thus not the case that research is developed outside the practice and that the field of practice in this way only serves as a place of application of research results (Brock-Utne, 1988, p.253).

The participative nature of the research undertaken was key in the research design. Students were involved in making decisions about the processes to be put in place and were actively engaged in evaluating the intervention they chose. Inquiry was completed *with* students rather than *on* students (Kemmis and Wilkinson, 1998).

Section 3: An exploration and critique of different action research models and an outline of the model adopted.

This section critiques two main action research models and concludes by presenting a retrospective model used to help guide the reader through the thesis.

Lewin (1946, 1951, 1952) is generally recognised as having introduced the concept of action research (see for example, Kolb, 1984; Kemmis & McTaggart, 1988b, Elliott, 1991 and Zuber-Skerritt, 1992). The model he proposed involved a "spiral of steps each of which is composed of a circle of planning, action and fact-finding about the result of the action" (1946, p.42). Later models of action research, for example those of Carr and Kemmis (1988) and Elliott (1991) are based on the work of Lewin, however there are two major problems with these models.

The first problem concerns the role and nature of reflection. For example, in Carr & Kemmis' (1988) model it is suggested that reflection is a discrete process when in reality it is present during the planning and acting stages. Indeed, Winter (1999, p.25) criticises the work of both Carr and Kemmis and Elliot (1991) for "the very cursory treatment given to the

process of reflection in two key texts on action-research which provide considerable detail on other aspects". Winter claims that reflection is the crucial process by which we make sense of evidence and that to ensure we learn from experience the process of reflecting must be given just as careful thought as the process of data gathering. "The fact that the process of reflection is largely taken for granted in these two texts suggests that reflection is tacitly assumed to be a straightforward process" (Winter, 1999, p.25).

Reflection, however, is *not* a straightforward process. It comes in many different forms that have different purposes. Cowan (1998, 2006), for example, differentiates between Kolb's model of reflection and that of Schön (1991). He describes Kolb's model of reflection as a component in a sequence, as "the bridge to be crossed between particular experience and consequent generalisation" and Schön's model of reflection as an "open-ended activity" (1998, p.37). Cowan's model shows the key role reflection plays in terms of prospective 'reflection-for-action'" contemporaneous 'reflection-in-action' and retrospective 'reflection-on-action' in any learning activity. Although Cowan does not discuss reflection in the context of action research, and despite Elliott and Carr and Kemmis' lack of discussion on reflection, it is clear that these forms of reflection are integrated into the planning and acting parts of each action research cycle. Consequently, any effective model of action research must reflect this.

A second problem relates to the purpose of action research models. An effective action research model would capture the process of action research effectively. Carr and Kemmis' (1988b) and Elliott's (1991) action research models are prospective models; they are *forward* looking. The elements of planning, acting, observing and reflecting appear to be discrete and follow neatly on from each other in both models as if it were possible to lay out in a diagrammatic format the plan for a number of research cycles *in advance*. Although they do acknowledge the unpredictable nature of social action and advise that it must incorporate a

degree of flexibility, there seems little room for manoeuvre in what appears to be a rather prescriptive, linear metaphor for the action research process. For example, the process appears to start with planning in Carr and Kemmis' model when in reality it is often reflection that results in the identification of a problem and prompts planning and action in the first place. The model suggests that reflection is a discrete process when in reality it is present during the planning and acting stages. The neatness of the model imposes a structure that does not reflect the actual messiness of the process in reality.

Referring to the work of Bourdieu, McNiff (2002, p.51) notes that there is often slippage between diagrams that communicate reality as sequential and predictable (for example those of Elliott, 1991 & Kemmis & McTaggart, 1988b) and the reality they purport to model. Such diagrams are produced as synopses of events. Bourdieu calls this a "synoptic illusion" because the model does not necessarily communicate people's experience of reality.

Burns (2000, p. 451) illuminates the main reason prospective action research models do not fit the bill: "action research does not know what questions to ask *until it has interpreted the present*" (Burns, 2000, p.451-452, my italics). It is clear from this that steps in an action research process cannot be designated in advance. It was only in *retrospect* that it was possible to construct a diagram of the process because at each stage of the research I could not foresee what questions I would need to ask until the present had been interpreted. The initial research question: 'what are students' perceptions of group work?' in Cycle 1 led to a second research question: 'what can be done to address the problem of social loafing in group work?' in Cycle 2 which could not have been foreseen until Cycle 1 was completed.

The model of action research used in the current study is loosely based on McNiff's work. McNiff (1988, 2002) proposes a visual metaphor that represents the process of action research which "is more appropriate to

the fluidity and unpredictability of practical living and the improvisatory knowledge base which underpins it" (2002, p.51/52). Her model is an upward spiral; "a spontaneous, self-recreating system of enquiry" (p.56) illustrated on page 9 of Chapter 1 with my research questions superimposed. Each of the three cycles of research carried out incorporates more than one iteration of planning and acting with reflection *for*, *in* and *on* action integrated rather than discrete from planning and acting. More information on the structure of each cycle can be found at Appendix 1.

Section 4: Methods

Research was carried out over a period of four years and included 729 first, second and third year undergraduates. It was undertaken using a mixed model design referred to as "intramethod mixing" by Johnson and Turner (2003, p. 298). This is defined as "the concurrent or sequential use of a single method that includes both qualitative and quantitative components", in this case questionnaires. They explain:

> The concurrent use of open- and closed-ended items on a single questionnaire and the sequential use of an open-ended questionnaire and a closed- ended questionnaire in a single research study are examples of intramethod mixing (Johnson and Turner, 2003, p. 298).

They argue that "in many cases, the mixing of qualitative and quantitative methods will result in the most accurate and complete depiction of the phenomenon under investigation" (p.299).

Intramethod mixing is defined as a pragmatic methodology (Tashakkori & Teddlie, 1998, p.19) that provides convergent and divergent evidence by combining the "complementary strengths and nonoverlapping weaknesses of quantitative and qualitative approaches" (Johnson and Turner, 2003, p. 298).

I decided to use questionnaires for two main reasons. It was important that the research was undertaken with a large sample to elicit the views and perceptions of a large number of students to produce warranted assertions for lines of action (Morgan, 2007). Questionnaires were the best tool to achieve this aim. The research was undertaken over four years with 794 students.

Secondly, this was real world research (Robson, 1993) with data provided by students during semester time. Some of the questionnaires were in the form of students' evaluations and data needed to be collected at the end of semesters which is the busiest time of the year for both students (completing assignments/revising for exams) and me (due to marking deadlines). Questionnaires given out and completed within sessions were deemed to be the most practical method for two reasons: they would ensure a good response rate and would take less time than interviews. Although interviews would have yielded more in-depth information about students' perceptions of group work, I felt that incorporating open questions into questionnaires would result in a breadth of viewpoints, not afforded by interviewing a smaller number of students.

Data Collection

Eight questionnaires were used over the four years of the research. During the first stage of research Questionnaire 1 aimed to ascertain students' perceptions of assessed group work so that I could identify which problems needed to be addressed. Having identified the problems and selected one key problem, Questionnaires 2 – 6 sought to establish students' responses to decisions that needed to be made before putting into action a process to address the problem identified in Stage 2 of the research. The final two questionnaires (7 and 8) were used to evaluate the intervention.

Questionnaires 1 and 8 were attached to module evaluation forms which were completed during the final session of the relevant module in order to achieve a good response rate. I was aware of the possible effects my dual role of "researcher-as-assessor" may have had on students. For this reason, students were given the chance to opt out of completing these questionnaires, were given the right to remain anonymous, and, signed consent was sought to use the data collected.

Table 3.1 shows the research question each questionnaire aimed to address, the number of students involved and the chapter in which findings of each questionnaire are discussed.

Chapter	Number of questionnaire and research question addressed	Class size	Sample size	Response rate
	Cycle 1: Initial research question			
4	1. What are students' perceptions of group work? What do they like about it and what problems do they experience?	147	107	73%
	Cycle 2: Choices made by students to inform decisions			
6	2. Would students prefer one group mark or should each student's mark be individualised based on peer assessment of contribution to group effort? Why? (Pilot)	34	31	91%
6	3. Would students prefer peer assessment of contribution to group effort, distribution of pool of marks or shared grade with exceptional tutor intervention? Why?	74	66	89%
8	4. Would students prefer criterion-referenced or holistic assessment?	99	92	93%
8	5. What % of marks should be allocated to peer assessment?	91	76	84%
8	6. What criteria should be used to assess peers' contribution?	135 (32 groups)	135	100%
	Cycle 3: Students' evaluation of intervention			
9	7. Was it good to be involved with setting criteria?	91 129	75 104	82% 81%
9	8. What are students' perceptions of summative peer assessment of contribution to group effort and of giving/receiving formative feedback from peers about their group working skills?	138	108	78%

Table 3.1 Outline of main questions addressed in each of the eight questionnaires used in the study.

My aim was to develop "a practical wisdom grounded in reflective experiences of concrete cases" (Elliott, 1991, p.53) in order to explore undergraduate students' perceptions of group work, identify the major problem(s) they reported, put in place an intervention to address one of these problems and evaluate students' responses to the intervention. The aim of the research was, ultimately, to improve my own teaching and learning strategies.

Ethical issues

Any intervention strategies designed to achieve this aim formed a natural and integral part of curriculum development. Students were not given the choice to participate in that the intervention would have been adopted in any case to try to improve their experience of group work and assessment as part of my normal practice. However, they were instrumental in choosing the nature of the intervention and in making decisions about how it was to be implemented.

Questionnaires were anonymous to protect students' identities although some students elected to include their names. Data analysis was carried out during the summer holidays after students' work was marked so there was no possibility of marks being affected by students' comments.

Qualitative data analysis

A number of the questionnaires included open questions on feelings that produced qualitative data (i.e. perceptions of group work in Questionnaire 1 and of formative and summative peer assessment in Questionnaire 8). Miles & Huberman (1994) describe qualitative data as an attractive nuisance while Marshall & Rossman (1999, p.151) note, "nearly as many analysis strategies exist as qualitative researchers." Mehmetoglu & Altinay (2006) claim that there is no single method of analysis in qualitative research (see also Dey, 1993; Denzin & Lincoln, 1994; and

Coffey & Atkinson, 1996). Robson (1993, p.370) concludes, "There is no clear and accepted set of conventions for analysis corresponding to those observed with quantitative analysis."

According to Patton (2002, p.453) content analysis refers to "any qualitative data reduction and sense-making effort that takes a volume of qualitative material and attempts to identify core consistencies and meanings." Content analysis was chosen as a method of analysing the qualitative data produced by open questions because, according to Burns (2000, p.432) and McKernan (1996, p.145) it is a useful means to identify hidden themes, concepts and meanings.

Also Robson (1993, p.371) suggests that open responses in questionnaires may be best dealt with using this method, which he describes as "codified common sense" (p.275). Using codified common sense appealed to my view of action research as a pragmatic approach.

The process involved two stages and adopted a grounded approach. First, open coding was carried out (Neuman, 2004; Strauss & Corbin, 1998). Creswell (2007) describes this as the process of developing categories. Recurring words or phrases were highlighted in a first attempt to ascertain students' feeling and perceptions from the text. Miles and Huberman (1984) refer to this as first-level coding. It involved the process of breaking down data and attaching descriptive labels (Willig, 2001) to units.

Secondly, the process of axial coding (Neuman, 2004) or second-level coding (Miles and Huberman, 1984) was undertaken. This is described by Creswell (2007) as the process of identifying interconnections between categories. It involved examining and reconsidering the initial highlighted units to look for concepts that clustered together. Whereas the previous stage of open coding involved breaking down the data; axial coding involved putting the data back together into holding pens which are described by Marshall & Rossman (1999, p.154) as "buckets or baskets

into which segments of text are placed." The aim was to create analytic labels (Willig, 2001) by making connections between categories and to organise the data into categories and their subcategories (Mehmetoglu & Altinay, 2006, p. 22, Creswell, 2007).

Quantitative data analysis

Questionnaires in Cycle 2 (Questionnaires 2- 6) involved mainly closed questions because students were asked to decide between various options (for example, criterion-referenced versus holistic peer assessment) so frequencies were calculated. Questionnaire 3 necessitated the formulation of a contingency table to show observed and expected frequencies for choices of assessment method. Questionnaires 1 and 8 involved calculating the frequencies of units in each category arrived at through the open and axial coding process described above together with the use of Cohen's Kappa or other methods to establish the effectiveness of intercoder reliability.

Ensuring rigour in quantitative and qualitative research

In order to ensure rigour five steps were taken.

1. Being explicit about the process of coding qualitative data

Questionnaire 1 was designed to ascertain students' perceptions of assessed group work. Data from this questionnaire were collected *before* the initial literature review on students' perceptions of group work had taken place using a grounded approach. The reason I chose to do this was so that I could approach analysis inductively and ensure (as far as possible) that emergent themes were grounded in the qualitative data and more context-sensitive i.e. that the themes were valid. Glaser & Strauss (1967) and McKernan (1996, p.7) stress the importance of allowing the

data to emerge on their own without any preconceived theories being imposed while Seidman (1998, p.100) warns:

> Most important is that reducing the data be done *inductively* rather than deductively. That is, the researcher cannot address the material with a set of hypotheses to test or with a theory developed in another context to which he or she wishes to match the data. (My italics.)

Miles and Huberman (1994, p.58) advocate a grounded approach on the basis that "data get well moulded to the codes that represent them, and we get more of a code-in-use flavour than the generic code-for-many-uses generated by a prefabricated start list."

However, Dey (1999, p.66) and Willig (2001) argue that categories do not simply 'emerge' from the data because they do not exist before the process of categorisation; rather they are suggested by the researcher during the research process. Nevertheless, it is possible to make a distinction between descriptive and analytic labels. Both are based on relations of similarity and differences but they function on different levels of abstraction (Dey, 1999, p.63). Descriptive labelling undertaken during open coding used indigenous (in vivo, or manifest coding) i.e. the actual words used by students were used as labels. It was only during axial coding that analytic labels were attached which involved exogenous coding. When exogenous coding was undertaken, samples of text were coded by a colleague to ensure intercoder reliability, i.e. the understanding I arrived at was corroborated by another source.

Morgan (2007, p.72) claims that in any case, there is no need to have to choose between the realist, objective, grounded, inductive approach preferred by Glaser & Strauss (1967), McKernan (1996) and Seidman (1998) and the subjective approach argued, for example, by Dey (1999) and Willing (2001). Instead, a pragmatic approach to conducting research suggests that in reality we work back and forward between a variety of frames of reference to establish an intersubjective understanding. It was important to try to triangulate frames of reference with a colleague to

arrive at an intersubjective understanding upon which warranted assertions could be made for particular actions.

Although in Questionnaire 1, open coding was conducted using a grounded (inductive) approach (Strauss & Corbin, 1998), using indigenous coding before the literature review into students' perceptions of group work had been undertaken, Questionnaire 8 used an abductive approach (Miles & Huberman, 1994; Morgan, 2007). Abduction involves the combination of inductive and deductive thinking; analysis involved making links between data and theories/ findings reported in the literature and data analysed inductively in Stage 1 of the study. In each case, a list of quotations which populated each category was compiled. Quotations were used extensively when presenting findings to increase the validity of the research.

2. Ensuring inter-rater reliability to warrant assertions made

Berelson (1952) warns, "since the categories contain the substance of the investigation, a content analysis can be no better than its system of categories" (in Robson, 1993, p.277). McKernan (1996, p.146) stresses the importance of establishing effective categories through which to filter the data and according to Silverman (2001, p.123):

> The crucial requirement is that the categories are sufficiently precise to enable different coders to arrive at the same results when the same body of material is being used.

Robson underlines the importance of testing the code on samples of text to assess reliability (p.279). He points out: "with a single observer, even if she shows high consistency, it may be that she is [carrying coding out] in a totally idiosyncratic fashion" (p.221). Robson (1993, p.221) defines inter-rater reliability as "the extent to which two or more observers obtain the same results when measuring the same behaviour" (e.g. when independently coding the same responses). It provides a credibility check

(Elliott et al, 1999) by confirming categories through convergence of different perspectives.

Inter-rater reliability involves the calculation either of the degree of correlation between the two sets of measurements (Tashakkori & Teddlie, 1998) or of the concordance between them (Robson, 1993, p.221). Miles and Huberman (1994, p.64) offer a more simple description: reliability equals the number of agreements divided by the total number of agreements and disagreements. Bakeman and Gottman (1986) advocate the use of concordance measures such as Cohen's Kappa, which correct for chance agreement (in Robson 1993, p.221). However Creswell (2007, p.211) follows Miles and Huberman's (1994) recommendation that calculating an 80% agreement is sufficient and suggests there is no need to calculate Cohen's Kappa. In order to accommodate both Creswell and Bakeman and Gottman's advice I asked different colleagues to code samples of data for each of the questionnaires which included qualitative data. Cohen's Kappa was measured when undertaking analysis of open questions in Questionnaire 1 while Creswell's method of establishing 80% agreement was used in the final cycle of research in Questionnaire 8.

Inter-rater reliability is important within Morgan's (2007) pragmatic approach to carrying out research. I was undertaking research to assess the workability of a potential line of action (in this case the implementation of a peer assessment intervention) based on warranted assertions (my interpretation of students' perceptions of the process). A pragmatic approach accepts that although we cannot hope for perfect understanding of others, calculations to measure inter-rater reliability can at least establish how much shared understanding of a data-set (in this case students' perceptions) exists. The higher the degree of shared understanding, the more effective the line of action designed to address issues was likely to be.

3. Supplementing qualitative data with quantitative data to establish transferability

Silverman claims that content analysis is reliable (if different coders arrive at the same results). However, he criticises content analysis because it can:

> furnish a 'powerful conceptual grid' (Atkinson, 1992, p.459) from which it is difficult to escape. While this grid is very helpful in organising the data analysis, it also deflects attention away from uncategorized activities (Silverman, 2001, p.123).

For this reason, an attempt was made to code all phrases (units). Units, which were not coded, numbered less than five per data-set and were assigned to a non-coded category.

Generalisability is a standard aim in quantitative research, achieved through statistical sampling procedures. There is some dispute, however, over the role and importance of generalisability in qualitative research. Some qualitative researchers (including Lincoln & Guba, 2000: Schwandt, 2000; Smith, 1984; Willig, 2001) argue that generalisability is not required in qualitative research. Stake (1994, p.236) however dismisses the 'intrinsic' study from which no attempt is made to generalise beyond a single case or even to build theories (in Silverman, 2001, p.249). I agree with Silverman (2001) and Mason (1996) who reject this idea. Mason notes:

> I do not think qualitative researchers should be satisfied with producing explanations which are idiosyncratic or particular to the limited empirical parameters of their study.... Qualitative research should produce explanations which are generalisable in some way, or which have a wider resonance (Mason, 1996, p.6).

However, Burns argues:

> An account can be judged to be externally valid if the *insights* it contains can be generalised beyond the situation(s) studied. .... It is only when the insights gained from a case study are *translated into an improved quality of action* that its external validity and

> therefore generality, can be demonstrated (my italics: Burns, 2000, p.450 - 451).

Improved quality of action is a key tenet of Morgan's (2007) pragmatic approach to research. To pragmatists the key question is 'what can we do with the knowledge gained from our research?' The external validity referred to by Burns (above) refers to applying the findings from such a situation to a second one that is similar (in Robson, 1993, p.405). Burns is, in fact, describing the action research process adopted in the current study. The "insights" revealed by analysing students' perceptions of assessed group work were designed to lead to "improved quality of action" in terms of their learning experience.

Morgan (2007, p.72) addresses the distinction between knowledge that is *either* specific and context-dependent *or* universal and generalised. He claims that we do not need to choose between extremes because research results are not likely to be so unique that they have no implications whatsoever for other practitioners in similar settings or so generalisable that they are applicable to all settings. The important thing in a pragmatic approach, according to Morgan, is that we take what we learn from the research and make the most appropriate use of that knowledge in other circumstances. 'Transferability' (Morgan, 2007, p.72) involves a solidly pragmatic focus on action (what can be done with the knowledge gained) rather than abstract arguments about the im/possibility of generalising the findings.

The methodological approach chosen was designed to possess what Burns (2000) refers to as "external validity" which Lincoln and Guba (1995), Henwood & Pigeon (1992) and Morgan (2007) call "transferability" and Silverman (1993) calls "generalisability" based on reflecting on the responses of a large enough sample of students to ensure assertions were warranted and action soundly based.

4. Guarding against anecdotalism and covert counting

Qualitative data analysis was supplemented by quantitative details of how often a comment was made for the following reason: to guard against anecdotalism and covert counting.

Coding and analysis of qualitative data has been criticised for being more of an art than a science (Burns, 2000, p.434, Robson, 1993, p.370) and for being guilty of anecdotalism which is described as the way some qualitative researchers use a few quotations from respondents to provide evidence of a particular contention (Silverman, 2001). Silverman raises doubt as to the persuasiveness of claims made on the basis of a few selected examples (p.35) while Bryman (1998, p.77) contends that:

> There are grounds for disquiet in that the representativeness or generality of these fragments is rarely addressed.

Mehan (1979, p.15) makes a similar complaint:

> Researchers seldom provide the criteria for grounds for including certain instances and not others. As a result, it is difficult to determine the typicality or representativeness of instances and findings generated from them.

Anecdotalism and selectivity undermine Dewey's pragmatic notion of 'warranted assertions' (Morgan, 2007; Johnson & Onwuegbuzie, 2004) because evidence is chosen rather than presented. For this reason qualitative data analysis was supplemented by quantitative details, for example, how often a word was used or a category assigned.

Robson (1993, p.401) differentiates between covert and overt use of numbers. He claims that qualitative researchers often use covert counting to make generalisations about trends or patterns in the data like "frequently", "rarely", and "commonly." Robson states that there is a strong case for the "overt and self-conscious use of frequencies" to generate actual numbers because the "necessary explicitness of

definition gives greater protection against bias" (Robson, 1993, p.401). Miles and Huberman (1994, p.41) state that quantitative data can help by showing the generality of specific observations and avoiding the "holistic fallacy" of monolithic judgments. For this reason the frequency comments were made was included in the report of findings.

Frequencies were calculated to substantiate the "trustworthiness" of my research (Johnson, 1999; Neuman, 2004; Lincoln & Guba, 1985). Miles and Huberman (1994, p.41) and Ecclestone (2008) suggest that quantitative data 'persuade' the reader and add more weight/credibility through de-emphasising the researcher's individual judgement. The element of quantitative data analysis was important since large samples of students were incorporated to guard against the criticism of "sacrificing measurement" (McKernan, 1996, p.7), and lack of reliability/generalisability, as a result of covert counting and the use of anecdotal evidence.

5. Following Henwood and Pidgeon (1992) and Elliott et al's (1999) guidelines for ensuring good practice in qualitative research.

Henwood and Pidgeon (1992) and Elliott et al (1999) suggest a number of guidelines for ensuring rigour in qualitative research. These include:

a) The importance of fit (Henwood & Pigeon, 1992) or grounding in examples (Elliott et al 1999). Analytic categories generated by the researcher should fit the data well. Quotations were used to allow the reader to assess the fit between the data and the researcher's interpretation of it.

b) Reflexivity (Henwood & Pidgeon, 1992) or owning one's perspective (Elliott et al, 1999) refers to the need for the researcher to disclose his/her own values and assumptions to allow readers to interpret the analysis and to consider possible alternative interpretations. Section 1 at the beginning of the

chapter explained my rationale for a pragmatic approach to methodology within a social constructivist epistemology, making clear the assumptions underpinning my approach to research.

c) Providing credibility checks (Elliott et al, 1999). I provided a credibility check by gauging others' interpretations of the data (through establishing inter-rater reliability).

d) Transferability (Henwood & Pidgeon, 1992) is referred to as situating the sample by Elliott et al (1999). To allow readers to explore the extent to which the study may or may not have relevance and applicability beyond, the specific context within which the data were generated was explained.

e) Accomplishing general versus specific research tasks. Elliott et al (1999) state that if the research seeks to develop a general understanding of a phenomenon, the researcher needs to ensure that the study is based upon an appropriate range of instances. Research was carried out with over 790 students over a period of four years.

Section 5: Summary

Action Research is an approach to improving education through implementing practical changes. The reason for the proposed research stemmed from a personal and professional need to meet the practical challenge of facilitating group work within a higher education setting. My aim was to improve the experience of group work for students and my own practice.

An action research approach was chosen for four main reasons: it offered a pragmatic approach to problem-solving; its aim is to improve practice; it

is participatory/collaborative and the cyclical nature of the action research process reflected the cyclical nature of the learning process it recorded.

It was participatory in that it involved me in my own inquiry and it was collaborative in that it involved students as part of a shared inquiry (Zuber-Skerritt, 1992; Brock-Utne, 1988; Winter, 1989; Whitehead & McNiff, 2006). Decisions for action were based on the evaluations of students involved in the research (Winter, 1989; Zuber-Skerritt, 1992, p.12 – 14; Berg, 2004, p.196; Whitehead & McNiff, 2006). It was research *with* rather than *on* students (McNiff, 1988, Creswell, 2007). It was practical, in that I was working as a change agent to improve my practice. The description of action research as a cycle involving planning, acting, observing and reflecting describes the natural process in which I was involved to improve my practice.

An action research approach accords well with a social constructivist epistemology which underpins my professional practice. Such an epistemology holds that knowledge production is not a phenomenon of an individual mind but is a social construction. The approach to research was based on Morgan's model of pragmatism. Action (the intervention to address the key problem with group work identified by students) was decided on the basis of warranted assertions with an emphasis on workability (Morgan, 2007). Warranted assertions were arrived at by establishing shared understanding of students' comments through an inter-rater reliability mechanism.

A pragmatist approach holds that qualitative and quantitative methods may be combined to address research problems. Questionnaires were the main method of data collection and, although usually associated with quantitative data, included open questions, which yielded qualitative data. The decision to gather both quantitative and qualitative data was influenced by the intent to examine processes and values associated with students' perceptions of group work and at the same time involve a large number of students in the process. The paper is written in the first person

to emphasise the connections between the writer and the material (Cooper, 2000, p.279).

Silverman (2001) claims that all research, irrespective of the paradigm from which it emanates, should be rigorous.

> If qualitative research is to be judged by whether it produces valid knowledge, then we should properly ask highly critical questions about any piece of research. And these questions should be no less probing and critical than we ask about any quantitative research study" (Silverman, 2001, p.221).

Although various theorists suggest that the concepts of validity, reliability and generalisability are inappropriate when assessing qualitative data (for example, Willig, 2001) and others claim that universal criteria against which qualitative research can be judged are still at an embryonic stage (Smith, 1996, Whitehead & McNiff, 2006), I believe that the concepts of reliability and validity are relevant to all research regardless of the paradigm from which action emanates. Reliability is the extent to which the results are replicable. By working with multiple groups of students using action research over a number of years rather than small groups of students on one occasion, and adopting measures to ensure inter-rater reliability, my aim was for my research to be reliable and generalisable. By using a grounded approach involving indigenous coding during the first stage of the research, and by furnishing quotations to support the exogenous approach to coding adopted in the final stage, I aimed to increase the validity of my findings.

Chapter 4: Students' perceptions of group work

This chapter addresses the first research questions in the preliminary cycle of action research: What are students' perceptions of group work; what do they like about it and what problems do they experience? Section 1 outlines the method. Section 2 reports the findings and discusses these findings in relation to the literature. It concludes by outlining the problems identified together with action to orientate the second cycle of research.

Section 1: Method

One hundred and forty-seven first year undergraduates nearing the end of a module on Human Communication in semester one were asked to complete a questionnaire as part of the module evaluation process. Part of the form consisted of two open questions on their experience of group work during the twelve-week course. Questionnaires were anonymous. The two open questions were:

- What did you like best about working as part of your group during this module?
- What did you like least about working as part of your group during this module?

Students generally worked in groups of four (although sometimes groups of three and five were unavoidable) on two assessed tasks:

- **A poster.** (30% of module marks.) The assessment criteria were generated and agreed by the students who, in groups, discussed each of the other groups' posters and provided formative written feedback and a summative mark. I provided a summative mark together with additional written feedback.

- **An essay.** (70% of module marks.) A semiotic analysis of a text. Individuals in each group agreed on the text, discussed it but produced individual essays. The criteria for assessment are dictated by the university but students had the opportunity to discuss and mark two essays from the previous year to ensure they understood the criteria.

A literature review was carried out after this first cycle of research had been undertaken because, when analysing students' perceptions of group work, I wanted the themes to be grounded in and emerge from the data as far as possible.

O'Brien uses the metaphor of a kaleidoscope to describe how we view our data:

> A kaleidoscope is a child's toy consisting of a tube, a number of lenses and fragments of translucent, coloured glass or plastic. When you turn the tube and look down the lens of the kaleidoscope, the shapes and colours visible at the bottom, change. As the tube is turned, different lenses come into play and the combinations of colour and shape shift from one pattern to another. In a similar way, ... by shifting theoretical perspective the world under investigation also changes shape (O'Brien in Silverman, 2001, p.293).

Although there is no "correct" view through the kaleidoscope, I wanted to be as objective as possible. Silverman, (2001, p.227) differentiates between "etic" and "emic" analysis. The former is based on the researcher's concepts whereas the latter derives from the conceptual framework of those being studied. I wanted, as far as possible, to adopt the latter approach. The discussion below compares my findings with the findings reported in the literature.

Section 2: Report and discussion of findings

Students' responses to the two questions outlined in the section above were analysed using content analysis. One hundred and seven students

completed questionnaires (response rate 73%). Each respondent's answers to the two open questions were divided into units that were coded using the method described in Chapter 3.

Cohen's Kappa was measured to calculate inter-rater reliability. The index of concordance was 96%. Cohen's Kappa was calculated as 0.957. Fliess (in Robson, 1993, p.223) suggests that Kappa of above 0.75 is "excellent".

Respondents made a total of 391 comments, 232 (59%) in answer to the question "What did you like best about working as part of your group during this module?" and 159 (41%) in answer to the question "What did you like least about working as part of your group during this module?"

The themes that emerged from these positive and negative comments are illustrated in Tables 4.1 and 4.2 respectively.

% of positive comments	Number of positive comments	Positive comments about the experience of group work on one module
32%	74	**The social aspect of group work**: including - meeting new people; meeting people from different cultures; making friends; having fun; a sense of community
23%	53	**Discussing/sharing of ideas**
21%	49	**Learning related aspects**: Including: learning people skills; learning from other group members
17%	40	**Workload related aspects**: sharing responsibility for work
7%	16	**Support:** receiving/giving support/help and encouragement
Total 100%	**Total** 232	

Table 4.1 Themes reflected in students' positive responses to group work

% of negative comments	Number of negative comments	Negative comments about the experience of group work on one module
30%	47	**Social loafers**: having to rely on others who did not pull their weight; sharing the same grade despite differences in individuals' contributions.
24%	38	**Organisational issues - meetings and time**: difficulty meeting due to different timetables, jobs or distances to travel; the extra time involved.
14%	23	**Differences of opinion**: not agreeing; conflict; confrontation; not having the same goals.
12%	19	**Perceived lack of skills**: not having the skills to cope with problems that occur in group work or with the different way of working demanded in group work.
9%	15	**Pressure**: not wanting to let others down; pressure due to the responsibility of being part of a group; not liking shared responsibility; the feeling of not being in control; feelings of "frustration" or "hassle"
6%	10	**Positive** comments in the negative section
2%	3	**Chatting**: too much chatting; going off at a tangent
2%	3	**Friends**: limited number of people worked with; not being able to choose whom to work with
1%	1	**Not coded**
Total 100%	**Total** 159	

Table 4.2 Themes reflected in negative comments about the experience of group work.

The social aspect of group work

Fifty-seven out of the 107 students (53%) who completed the questionnaire mentioned the social aspect as one of the aspects they liked best about their experience of working as part of a group during the module. The social aspect of group work accounted for 74 out of the 232 positive comments (32%) made by respondents. Since this module runs in semester 1 of Year 1, respondents appreciated the opportunity to get to know other students at the beginning of their university career. For example:

> "It helps you settle in in the first term." (Respondent 18)
> "I thought it was a good module to have as your first as it made you meet people within your class." (Respondent 65)

"Getting to know one another; getting to know people in the same halls as me." (Respondent 88)

Research by others resulted in similar findings. Students in studies by Garvin & Butcher (1995); Stanier (1997); Agyemang & Unerman (1998); Maguire & Edmondson (2001) and Strachan and Wilcox (1996), for example, also reported that they enjoyed the social aspect of learning in a group.

Students in the current study also reported that they liked the opportunity to make new friends and get to know people from different backgrounds:

"I believe it was a good way to make friends." (Respondent 41)
"You got to really know the members of your group and make new friends." (Respondent 66)
"It was good to get to know people I probably wouldn't have met otherwise." (Respondent 23)

Similarly, Stanier's (1997) research with 35 second-year undergraduates found that 25 out of 35 reported benefiting from acquiring a different network of friends as a result of group work.

Students in my study reported that the social aspect made learning more enjoyable:

"It made it more enjoyable working in a group." (Respondent 29)
"The social aspect; it was good fun meeting as a group to do work." (Respondent 36)
"Social side; was able to work well together but also have fun." (Respondent 72)
"It also made the work feel less like work!" (Respondent 78)
"Provides a social aspect to work and makes assignments a more enjoyable experience." (Respondent 82)
"The friendliness and relaxed working environment." (Respondent 90)
"It made the work far more interesting." (Respondent 97)

This confirmed the findings of others, for example Falchikov (1993); Levin (2005) and Cheng & Warren (2000).

Students also referred to a sense of belonging, feeling part of a community:

> "It's nice to feel part of a team." (Respondent 40)
> "No one feels left out." (Respondent 53)
> "Sense of community." (Respondent 59)
> "Didn't feel so alone and overwhelmed." (Respondent 64)

Discussing; sharing ideas

The interaction that resulted from group work was seen as both an advantage and a disadvantage in my research and also in research carried out by Strachan & Wilcox (1996); Bourner et al (2001); Phipps et al (2001); Morgan (2003); Barfield (2003); and Knight (2004). Students appreciated the support from other group members but, on the other hand, did not like having to rely on others since often the latter proved to be unreliable. Thirty per cent of positive comments in my research related to the advantages of interaction such as the support and help available from peers together with the opportunity to discuss and share ideas. This accords with the findings of Doise and Mugny (1984); Bennett (1991); and Webb (1995).

Forty-one respondents in my study (38%) made 53 comments (23% of all positive comments) reflecting students' liking for discussion and being able to share ideas. For example:

> "Being able to discuss things." (Respondent 6)
> "The brainstorming sessions we did. I enjoyed bouncing ideas off one another." (Respondent 8)
> "I liked the ability to sit and talk through individual difficulties with the material." (Respondent 57)
> " We could share our ideas. " (Respondent 71)
> "It was good to hear other people's opinions." (Respondent 86)

Students also appreciated that there was more opportunity for discussion than lectures or large teaching groups allowed and that the smaller group made them feel more comfortable voicing opinions:

"We heard everyone's opinion." (Respondent 47)
"Everyone has their say." (Respondent 53)
"We have more opportunity to speak." (Respondent 54)
"It's easier to talk and ask something than in a big group." (Respondent 105)
"The group environment is more comfortable to share your ideas than the whole class group." (Respondent 65)
"Being able to ask about things I wouldn't feel comfortable asking the lecturer about." (Respondent 88)

Learning related aspects

Forty-three students made 49 (21%) comments related to learning. Students reported liking the opportunity to acquire people skills, for example:

"Working in company and learning how other people work and organise themselves." (Respondent 3)
"You get to learn about different cultures." (Respondent 7)
"Learn how to listen to others' opinions and be firm about things when you have to be." (Respondent 13)
"It also gave me the chance to improve my communication skills." (Respondent 20)
"Learned about group communication and how to work together." (Respondent 42)
"It helped me learn how to communicate effectively and how to listen to other people's views." (Respondent 95)
"Learning how to solve problems that arise in group conflict." (Respondent 98)
"I was able to learn how difficult it is to work as part of a group; after each meeting and each group assignment I learned what is necessary for group work – communication skills, personal qualities etc; it was challenging but I liked learning it." (Respondent 103)
"It gave an opportunity to learn more about people's psychology." (Respondent 104)

Students also liked learning from other members of the group:

"You get better results by inspiring each other with your own ideas – gets you to think more creatively about your assignment." (Respondent 7)
"Working in a group gave me a lot of good ideas which I might not have had." (Respondent 20)

"Being able to discuss about the module which I found very helpful and useful; it helped me to understand more about my module and what I'd learnt." (Respondent 83)

These findings confirmed research undertaken by others (for example Bennett, 1991, Webb, 1989,1991,1995) which found that it is the interactions that occur in groups that facilitate learning.

Workload related aspects

Sharing workload was seen as a positive aspect by students in my research (see also Maguire & Edmondson, 2001 and Knight, 2004) but was also seen as negative due to the lack of control it brought with it (see also Maguire & Edmondson, 2001) but only two students admitted to feeling pressurised by "the responsibility" (Respondent 20) and "the pressure and the feeling that I would be letting people down if I don't do brilliantly" (Respondent 19). Students in Morgan's (2003) study also identified this feeling of pressure. Pressure to meet group goals and make a valuable contribution to a team exists in the world of work, so it is important that students experience this. The individualistic nature of the traditional path through university often denies students this experience.

Forty comments (17% of all positive comments made) related to workload related aspects. Thirty-eight out of the 107 respondents (36%) gave workload related answers to the question: What did you like best about working as part of your group during this module?

Students liked being able to share workload and the responsibility for it:

"Each taking on different roles and responsibilities." (Respondent 39)
"And it made the workload lighter as you could share research." (Respondent 36)
"That it takes some of the pressure off you individually as four heads are better than one." (Respondent 37)
"Sharing the responsibility for the work." (Respondent 56)

"I liked best the shared responsibility for our assessment." (Respondent 57)

They also liked being able to play to their strengths to produce a piece of work:

"Everybody contributes different aspects." (Respondent 17)
"We got to know each other's strengths and weaknesses and therefore could work together efficiently." (Respondent 43)
"Everyone had the opportunity to show and incorporate their own talents; any weaknesses were covered by other group members' strengths." (Respondent 80)

Responses by Respondents 43 and 80, above, reflect the tension between using individuals' strengths and abilities to achieve a better product and enabling group members to develop skills in areas which they most need to develop, highlighted by McDowell (1995, p.311).

Receiving/giving support and encouragement

Students also appreciated the support and encouragement of other group members. Fourteen students (13%) made a total of 16 comments (7% of all positive comments made) referring to this aspect. For example:

"If I did not understand something, they were there to help me." (Respondent 49)
"You could get guidance from group members if you were doing things wrong." (Respondent 52)
"The help .... and reassurance." (Respondent 59)
"Got support from the other group members if needed and gave support." (Respondent 64)
"When one of the group members was suffering from assignments, others could help and explain." (Respondent 102)

Organisational difficulties

Students were also asked what they liked least about working as part of their group during the module. The 107 respondents' answers resulted in 159 negative comments about their group work experience.

There were very few downsides to the social aspect identified by my students other than the organisational difficulties associated with having to meet up (also identified by students in the study by Maguire & Edmondson, 2001) and the tendency identified by three students to chat rather than focus on work.

However, one of the aspects students liked least about their group work experience was the difficulty of meeting as a group outside taught sessions. The problem of coordinating such meetings (often around part-time jobs) exists in all university settings but is exacerbated by two factors at the university in which the research was undertaken. The first problem lies in the scale and flexibility of the modular degree structure which means that first year students (once they have registered for the compulsory modules specific to their single or combined degree course) are free to choose any other level 4 modules they like. Modules run in 'slots' timetabled in the morning, afternoon or evening so evening meetings are not always possible. Students on first year modules, as a result, can be from virtually any field of study, each individual with a different timetable.

The second problem arises from the fact that there are three campuses in different parts of the city with first year student accommodation at each campus and additional accommodation elsewhere. Travel between each campus is not straightforward and can be very time consuming. Road works and congestion often exacerbate the problem.

Twenty-six students cited difficulty meeting up as one of the aspects they least liked (24%). For example:

"We all had different timetables to work around so it was not always easy to meet." (Respondent 96)
"Having to find a time when everyone was available. None of us lived near each other which also didn't help." (Respondent 25)

Time was problematic in other ways too and was considered an issue by nine students. Respondents 97 and 7 commented respectively, for example, that it "took longer to complete tasks" and "it takes so much more time." Three students commented on other organisational issues such as "always having to fit work in around the group" (Respondent 74). This is likely to be more of a problem in the second semester and in years 2 and 3 when students are more likely to have jobs.

Differences of opinion

Fourteen per cent of all negative comments (n = 23) related to some sort of disagreement. This included people not listening. For example Respondent 77 complained: "Some members didn't listen or accept criticism very well" while Respondent 44 disliked " people not listening to your ideas" and Respondent 48 complained "and no matter what you say, they won't listen." Differences of opinion were also problematic e.g. Respondent 13: "At times there can be differences of opinion" or "confrontation" (Respondent 34), "or when conflict arises" (Respondent 26). There were also clashes of personality "if one group member is quite bossy and not flexible" (Respondent 32) and "some people can dominate groups" according to Respondent 65.

Livingstone & Lynch (2000) also found that negative comments from students tended to focus on conflicts of interest, which had a detrimental effect on group functioning. One interpretation of this is that students tend to view conflict affectively, as a negative consequence of interaction rather than as a cognitive tool to improve skills, knowledge and understanding. However, Livingstone & Lynch (2000, p.340) also noticed that in groups where roles were defined, personality clashes were

moderated and completion of the task was prioritised while clashes were shelved. An implication of this interpretation is that increased use of roles and knowledge of group functioning would be beneficial.

Perceived lack of skills

Phipps et al (2001) found that students did not know how to work in groups unless shown. Twelve per cent of students in my study (n = 19) acknowledged a lack of skills or experience to be able to cope with problems that occur when working as part of a group. For example, Respondent 7 commented: "You really need to be patient"; while Respondent 99 concluded, "sometimes it was too difficult to solve problems." Perceptions of deficits in communication skills such as negotiating as well as decision-making skills were apparent, for example Respondent 4 disliked struggling to "come to a compromise" and Respondent 11 acknowledged: "it's a bit difficult to come to a conclusion...especially on how the work is to be done." Respondent 2 disliked "people not being able to adjust their style of working to accommodate to working as part of a team."

Social loafers

The most cited complaint (articulated by 30% of respondents) was the perceived lack of effort by one or more members of the group which can be summarised as: "When group members weren't pulling their weight" (Respondent 55). Students specifically disliked having to share the same grade with peers who did not contribute equally:

> Respondent 28: "If one member does not perform, all will lose marks as a result."
> Respondent 33: "All achieving the same grade for assignments even though some members contributed more and turned up regularly for group meetings, while others didn't."

> Respondent 42: "some people did slightly more work than others but still all got the same grade."
> Respondent 52: "If people didn't work hard then your grade might not be as good as it could have been."
> Respondent 86: "Sometimes members of the group relied on others doing the work and the work load was less evenly spread but everyone got the same mark."

In the studies located in the literature the issue of students who did not pull their weight was of key concern. These students were variously called "hitchhikers" (Mesch, 1991), "slackers" (Colbeck et al, 2000), "freeriders," (Kerr & Bruun, 1983; Cheng & Warren, 2000), "freeloaders" (Underwood, 2003) "social loafers," (Latané et al, 1979; Harkins, 1987; Karau & Kipling, 1993; Sheppard & Taylor, 1999; Gagne & Zuckerman, 1999; Underwood, 2003: Falchikov, 2005) "passengers" (Stanier, 1997; Bourner et al, 2001) or " coasters" (Ross & Rolheiser, 2003).

Social loafing or freeloading, etc. is also described as the "tendency of individuals to reduce the effort expended towards a task when working in a group, resulting in a disproportionate burden of responsibility on the willing or active members of the group" (Underwood, 2003, p.331). According to Underwood (2003) antipathy to group work is often associated with the experience of social loafers or freeloaders whom she defines as "individuals who withhold effort if they can achieve their educational goals by letting others do the work" (p.321). (See also Walker 2001.) For example, responses by students in Bourner et al's (2001, p.24) study raised the issue of "passengers" and 12 of the 35 students in Stanier's (1997) study complained about the issue of "passengers". Students in Livingstone & Lynch's (2000, p.339) study complained about students not turning up for meetings and not doing what they were asked to do. Gillies and Ashman (2003, p. 69) identified students who did not pull their weight as a major student concern.

Conway & Kember (1993, p.50) introduced peer assessment of an individual's contribution to a group project as a result of unsolicited student comments on the iniquity of awarding the same mark to all

members of a project group. Students in Phipps et al's (2001) study also complained about shared grades. Students in the former study had reported that they found group projects "more effective than lectures for most aspects of learning" (Conway & Kember, 1993, p. 45) but were concerned about the fairness of assessment when all group members were given the same mark in spite of different levels of contribution. Students in my study also questioned the fairness of receiving a group mark when not all members had contributed equally.

Students in my research perceived lack of reliability and social loafers as the most negative aspect of group work. Social loafing was the dislike most often cited by students. However, only one respondent (Respondent 36) admitted to falling into this category:

> Although you could use the work of others which reduced the work, it made me lazy and I didn't put as much effort in as I would if I was responsible for all my work.

There are at least two possible interpretations of this: students did not see themselves as social loafers or they only saw social loafing as a disadvantage if it was someone other than themselves doing the loafing. Karau & Kipling (1993, p. 696) found that participants in their research were either not aware that they engaged in social loafing or were unwilling to report that they engaged in social loafing.

Assessment of group work was seen as a key issue not only by students but also by tutors involved in other studies located in the literature. Mello (1993, p.254) for example, identified two main problems with group work: conflict among members and "those individuals who do not do their share of the work but reap the benefits of their more productive group members." Cheng & Warren (2000, p.245) claim that "the kind of free-riding problem or unfairness of uniform grading within a group ..... can be discouraged by awarding students marks or grades based on their individual contributions to the group work."

Assessment of individual contribution to the group product is called for by many researchers with experience of group work (for example, Falchikov, 1988, 1991, 2005; Goldfinch & Raeside, 1990; Conway & Kember, 1993; Butcher & Stefani, 1995; Lejk & Wyvill, 1996; Strachan & Wilcox, 1996; Rafiq & Fullerton, 1996; Krause & Popovich, 1996; Pond & ul-Haq, 1997; Healey, 1999; Cheng & Warren, 2000; Li, 2001; McWhaw et al, 2003; Kilic & Cakan, 2006; Sharp, 2006, Bloxham & Boyd, 2007). Kagin (cited in Morgan 2003, p.2) goes as far as claiming, "every time I see group grades being used I am appalled. They are, in my view, never justified. Ever. " According to Sharp (2006) "The best measure of the contribution of each student (to a group project) is the sum of the other students' evaluations of him/her" (p.332).

Underwood (2003) found that students were very reluctant to take any overt punitive action against non-contributors – four out of five students in her study said they would take no action. Livingstone and Lynch (2000, p.340) also identified reluctance to take action against social loafers. If this is the case, then the introduction of structured summative peer assessment of contribution to the group effort would be one solution.

There are many arguments in favour of introducing a system to differentiate between each group member's contributions to an assessed group product. Krause and Popovich (1996, p.143) assert: "A system of accountability must be implemented that rewards those who contribute and penalises those who do not." Since group members are in a better position to judge individual contributions to the group's performance than the tutor, there is a strong argument that this function should be performed by the group members themselves (Magin, 2001; Bushell, 2006, p.91). However Lejk & Wyvill (1996, p.267) acknowledge, "converting the effectiveness of an individual's contribution to a group into a numeric grade is a complicated and problematic task " and Ross & Rolheiser (2003, p.120) point out the difficulty of having to "disentangle individual from collective performances because students who coast on the work of others must be identified."

Summary

This first cycle of action research found that students had more positive than negative perceptions of group work. They particularly enjoyed the social aspect, which involved meeting new people, making new friends, making learning fun and developing a sense of community. They also liked the process of discussing and sharing ideas with others and learning from their peers as well as learning people skills. They appreciated giving and receiving support from peers and sharing the responsibility for work. However, a quarter of all negative comments related to organisational issues such as conflicting timetables and the difficulty of travelling between campuses. Also some students believed that they did not have the skills needed to manage group work effectively and viewed conflict as negative. However, students' major concern about group work related to peers who do not pull their weight. Underwood (2003, p.333) refers to such students as "parasitic peers." However, these students are generally referred to as social loafers, freeriders, coasters or passengers in the literature. For the purposes of this study, the term "social loafer" will be adopted.

Action plan undertaken to address problems identified

1. Discussion facilitated to make students aware of the benefits of conflicting viewpoints.
2. Skills training provided in an Interpersonal Skills module. A module on Communicating in Groups is also provided in semester 2.
3. Students divided into groups based on where they live and when they have free time.
4. A summative peer assessment of contribution to group effort will be piloted next year.

Chapter 5: Methods designed to alleviate social loafing

In Chapter 4 the issue of social loafing was identified as a key problem in group work by students on a first year module. A review of the literature confirmed that the problem was widespread. Section 1 explores the theoretical explanations for social loafing discussed in the literature. It then summarises the findings from the literature review and suggests implications for practice, including the introduction of a peer assessment mechanism. Section 2 begins with a definition of peer assessment. It offers a brief overview of the methods available before considering the three main methods in detail and discussing students' responses in the literature. It concludes with a discussion of the implications for practice.

Section 1: Theoretical explanations for social loafing

For some sorts of tasks the presence of others results in improved individual performance so groups are frequently used to enhance productivity. This phenomenon is referred to as social facilitation (Zajonc, 1965). However, working as part of a group can also have a deleterious effect on individual productivity. Since collective work settings are so pervasive and indispensable, research identifying which factors motivate and demotivate individuals within these collective contexts together with strategies to overcome demotivation is clearly needed (Latané et al 1979, p. 84; Karau & Williams, 1993, p.681).

In an experiment undertaken in 1913, Ringelmann found that when individuals within a group perceived that they could neither receive their fair share of rewards nor the appropriate blame, they frequently reduced their effort. Participants were asked to pull on a rope attached to a strain gauge with as much effort as possible. Ringelmann found that group productivity failed to reach the levels predicted on the basis of individual performance. Dyads performed at 93% of their potential average ability, groups of three performed at 85% of their potential average ability, and

groups of eight performed at 49% of their potential average ability (in Schnake, 1991, p.42).

The Ringelmann Effect (later referred to as social loafing) thus describes the inverse relationship between the size of a group and the magnitude of a group member's individual contribution to the accomplishment of the task (LaFasto & Larson, 2001, p. 77).

The term 'social loafing' was originally coined by Latané et al (1979) to describe the reduction in effort of people working collectively as opposed to coactively, ascribing it the status of a social disease. (Working collectively is defined as "when individuals work in the real or imagined presence of others with whom they combine their inputs to form a single group product" while working coactively is defined as "when individuals work in the real or imagined presence of others but their inputs are not combined with the inputs of others" (Karau & Kipling, 1993, p.681)). Between 1974 and 1993, over 80 studies on social loafing were carried out in which individuals' coactive efforts were compared with individuals' collective efforts. According to a meta-analysis of these studies carried out by Karau and Kipling (1993) social loafing is evident across a wide variety of tasks including physical tasks such as rope pulling and cycling; cognitive tasks such as identifying radar signals and navigating mazes; creative tasks such as brainstorming and song writing; and, evaluative tasks such as assessing the quality of poems and essays. Despite the vast majority of research having been undertaken in the context of a 30 – 60 minute experiment, according to Karau & Williams (1995, p.135) social loafing generalises across tasks as well as most subject populations.

It is possible to synthesise explanations for social loafing into seven categories which are discussed overleaf.

Submaximal goal setting

Latané et al (1979, p.829) suggest that group members might expect it to be easier to achieve a goal when others are helping and reduce their effort as a result.

Matching of effort

Another explanation for social loafing involves matching of effort. This position suggests that social loafing occurs because individuals expect or think they see other group members reduce their efforts when working in groups and, as a result, reduce their own efforts to maintain equity. Group members see no reason to work hard to compensate for "shirkers" (Latané et al, 1979, p.829). Matching of effort also helps alleviate what Schnake refers to as the "sucker effect" (1991, p.42). Since group members tend to want to avoid being a "sucker" they decide to withhold effort themselves. This research is important because it suggests that establishing high expectations as a group norm is an important part of group processing.

Identifiability

Another interpretation of social loafing involves the concepts of identifiability and evaluation potential (for example, Latané et al, 1979, p.829; Karau & Williams, 1993, p.683; Price & Harrison, 2006, p.1375). Kerr and Bruun (1983) for example argue that social loafing occurs because in most groups input is combined into one group product and individual contributions are hidden as a result. Individual members can "hide in the crowd" (Latané et al, 1979, p. 830). Since they cannot be singled out for credit or blame, their incentive to perform is reduced and they engage in social loafing (Harkins and Petty, 1982).

Latané et al (1979), Williams et al (1991), Harkins (1987) and Karau & Williams (1993) suggest that making individuals' collective inputs identifiable may be enough to eliminate social loafing in many situations. However, Harkins & Jackson (1985) found that although, when outputs were individually identifiable, participants were more productive than when their outputs were pooled, this difference emerged only when participants believed that their individual outputs could be evaluated through comparison with other group members' performances (evaluation potential). When participants believed that their individual outputs were not comparable and thus could not be evaluated, there was no difference in performance by participants whose outputs were identifiable and those whose outputs were pooled. Harkins and Szymanski (1989) also found that people were less likely to loaf on collective tasks (even though individual outcomes could not be evaluated) if they believe the performance of their group is being compared with the performance of other groups. This research suggests that in order to eliminate social loafing group members must not only feel that their outputs are individually identifiable as suggested by Williams et al (1981) but that these outputs can be evaluated through comparison with the outputs of other in-group members, and that the group's overall performance is being compared with that of other groups (Harkins & Szymanski, 1989). There must, therefore be a standard (personal, social or objective) with which these outputs can be compared (Karau & Williams, 1995, p.136).

This research is useful because it demonstrates the importance of each group member's output being identifiable and comparable and the motivation potential of inter-group competition. A peer assessment system would fulfil the need for identifiability; if the students were involved in setting the criteria, standards would be clear and comparability facilitated; and if groups assessed each other's outputs, inter-group competition would be established.

Dispensability

Another possible cause of social loafing involves dispensability of effort. Kerr and Bruun (1983) found that individuals tend to exert less effort when working as part of a group if they feel that their contribution is not essential to a high-quality group product i.e. if they feel their contribution is dispensable. Hardy & Crace (1991) found that when members of a team believe their ability dictates that they will make a limited contribution to a team effort, social loafing increases. Harkins & Petty (1982) and Sheppard & Taylor (1999) also found that individuals worked just as hard collectively as coactively when their individual inputs to the collective product were unique but loafed when their inputs were either potentially or completely redundant.

Steiner (1972) differentiates between disjunctive and conjunctive tasks. In disjunctive tasks, the group must decide on one contribution as the group answer (Meyers, 1997). As a result, disjunctive tasks allow success to result from the contribution of only one member (usually the brightest) for example solving a mathematical problem or an anagram. Conjunctive tasks, on the other hand, mean that the group product involves contributions by all members and can only be as good as the weakest member, for example, a relay race (Meyers, 1997) or a group climbing a mountain linked by ropes which can only go as fast as the slowest person (Kerr & Bruun, 1983, p.80).

Kerr and Bruun (1983) predicted that the effect of member ability on perceived dispensability would have opposite effects for conjunctive and disjunctive tasks. On a disjunctive task where only the best member's contribution mattered, they found the less able members social loafed because they saw their contributions as dispensable. On a conjunctive task, since only the least able member's performance mattered, the high ability members reduced their efforts because they were dispensable. In additive tasks, however, the group product is equal to the sum of each

group member's contribution, so Meyers (1997) suggests task structure should be additive to ensure group members feel indispensable.

Assigning clear roles for individuals in the group can ensure the contributions of each group member are indispensable. However, Kerr and Stanfel (1993) reported that when one group member was designated as a token leader (i.e. having a title of leader without any power, legitimacy or formal responsibility) the remaining group members showed decreased personal responsibility for group performance and higher rates of social loafing. Sharan and Sharan's (1976) suggestion that a leadership role should be established and then rotated would solve this problem.

It is also evident that group members are more likely to care about the effects of their indispensable actions if they value the group to which they belong. This underlines the importance of developing a sense of group cohesion. This can be achieved through what Johnson and Johnson (1991, p.89 & 2003, p.111) refer to as "promotive interaction" which can be achieved through giving groups time to get to know each other and by using icebreakers and intergroup competitions.

Research on dispensability is useful for three reasons: firstly, it suggests that the type of task and how it is structured is key to whether or not group members social loaf; secondly, it demonstrates that identifiability on its own will not alleviate social loafing in all cases because when individuals feel dispensable they will reduce effort even when their contribution is made identifiable to others involved (Karau & Kipling, 1993, p.683) and; thirdly, it underscores the importance of promotive interaction as a tool to establish group cohesion.

Expectancy-value theory

According to Sheppard & Taylor (1999) expectancy-value theory consists of three components:

Expectancy. The expectancy component refers to the perception that performance is contingent on effort (i.e. greater efforts will result in better performance).

Instrumentality. The instrumentality component refers to the perception that the consequence of the performance outcome is contingent on performance (i.e. that performance will determine the outcome). A student might believe that a good project will receive a good grade and a weak project will receive a low grade (high instrumentality), or she may believe that the teacher does not mark fairly and that there is, therefore, no relationship between the quality of the project and the grade allocated (low instrumentality) (Sheppard & Taylor, 1999, p.1147).

Value. Karau & Williams (1993, p.696) discovered that the tendency to engage in social loafing decreased as perception of the value of the task increased and conclude that high levels of task meaningfulness might eliminate social loafing. Sheppard & Taylor (1999, p.1148) refer to this as the value component i.e. how much value or importance the student attaches to achieving the outcome of the performance. In the case of a group project, the student's module mark may depend largely on the mark received for the group project thus making the mark important (high value). Or the group project mark may have a negligible impact on the student's overall module mark (low value). It is important to note two things. Firstly, value does not depend only on how important or rewarding the student views the outcome, it also depends on the psychological and material costs associated with achieving the outcome. So value refers more properly to how much value or importance the student attaches to achieving the outcome of the performance minus any costs incurred.

Secondly, rewards need not be extrinsic (for example, marks) if students are intrinsically motivated.

Effort motivation reflects how much effort a student is willing to put into a task and can be expressed as the product of expectation, instrumentality and outcome value, discussed above. In group work, according to expectancy-value theory, effort motivation should be high when students:

- Perceive a contingency between their effort and performance.
- Perceive a contingency between performance and the outcome.
- Value the outcome (i.e. the benefits associated with contributing or achieving the outcome must exceed the cost of contributing).

Social impact theory

Social loafing can also be explained in terms of social impact theory. Latané's (1981) social impact theory suggests that people can be viewed as either targets or sources of social impact. When people work as a group, the demands of an outside source of social influence (in this case, me as tutor) are diffused across multiple targets (i.e. all the group members), leading to decreased levels of effort. The amount of social impact experienced is related to the strength, immediacy and number of sources and targets present in a situation and is predicted to follow an inverse power function specifying that each additional group member will have less influence as group size increases. In a typical social loafing experiment, the researcher is a single source of social impact and the group members are multiple targets. Social impact theory suggests that the researcher's request to try as hard as possible on the task should be divided across targets (the group members) resulting in reduced effort as group size increases. When individual targets are not part of a group, however, they feel the full impact of the researcher's request and try harder (Karau & Kipling, 1993, p.683). This theory is useful because it underlines the importance of limiting group size. North et al (2000) for

example found that social loafing occurred less in groups of three than in groups of eight, confirming the importance of social impact theory. (North et al's study was replicated with the students on one module and their findings were repeated.)

Gender and cultural differences

Sheppard & Taylor (1999) found gender and cultural differences in terms of social loafing. They found that females were less likely to social loaf than males and that all-male groups were more susceptible to social loafing. In addition, individuals in Eastern cultures were less likely to engage in social loafing than individuals in Western cultures. They also found that:

- Social loafing was robust across tasks demanding different types of effort (cognitive, physical, perpetual).
- Social loafing occurred more when a deceptive cover story was used i.e. individuals were more likely to loaf if they were not told that the purpose of the research was to examine effort.
- Social loafing was eliminated when participants worked in highly valued groups.
- Participants either were not aware that they engaged in social loafing or were unwilling to report that they engaged in social loafing.

Summary

A major criticism of the research on social loafing is that most studies have been carried out with students in the context of a 30 – 60 minute experiment based on relatively meaningless tasks devised as experiments rather than in real-life settings. It could be argued that such experiments are ecologically invalid in that they do not draw upon real

teaching situations (Nicol & Macfarlane-Dick, 2007). Despite this, it appears from the literature that there are a number of variables likely to affect the preponderance to social loaf:

- **Group size.** When groups are too big (Karau & Kipling, 1993; Johnson & Johnson, 2003; Kerr & Bruun, 1993; Karau & Williams, 1995).
- **Identifiability and evaluation potential.** When group members' individual contributions are not identifiable or evaluable against a clear standard (Latané et al, 1979, p.829; Williams et al, 1981; Harkins & Jackson; Harkins & Szymanski, 1989; Slavin, 1989b, p.52; Karau & Kipling, 1993, p. 700; Karau & Williams, 1995, p.136; Gagne & Zuckerman, 1999, p.525; Gillies & Ashman, 2003, p.74; Johnson & Johnson, 2003; Price & Harrison 2006).
- **Dispensability.** When contributors' input to the collective outcome is perceived to be dispensable or redundant or when roles are not clearly defined (Harkins & Petty, 1982, Karau & Kipling, 1993, p. 700; Karau & Kipling, 1995, p.136; Kerr & Bruun, 1983, p.77, Hardy & Crace, 1991, p.379).
- **Task.** When a task is perceived as lacking challenge or low in value/meaningfulness (Harkins & Petty, 1982; Karau & Kipling, 1993, p. 700; Sheppard & Taylor, 1999, p.1148).
- **Promotive interaction.** Lack of group cohesion or when working with strangers or people who are not valued (Karau & Kipling, 1993, p. 700; Sheppard & Taylor, 1999, Johnson & Johnson, 2003).
- **Gender.** All male groups tend to social loaf more than mixed or all female groups (Kugihara, 1999; Sheppard & Taylor, 1999).

Implications for practice

The most common explanation for social loafing is that an inability to identify or evaluate individual performance in a group situation leads

individuals to conclude that nothing can be gained from effort spent on their performance (Williams et al, 1981; Gagne & Zuckerman, 1999). Although peer assessment of contribution to group effort may go some way to alleviating social loafing, there are other factors that need to be taken into consideration:

- Assigning meaningful tasks.
- Making sure groups are not too big.
- Enhancing cohesiveness of groups through promotive interaction so that students value each other.
- Enhancing group-processing skills.
- Building in opportunities for self-evaluation against a clear standard.
- Ensuring that the tasks assigned to individual group members are designed to be indispensable.
- Introducing inter-group competition.
- Avoiding all male groups.

Section 2: Peer assessment

In Section 1 a review of the literature on social loafing suggested that introducing summative peer assessment of contribution to group work was a key factor in alleviating social loafing. This section begins by defining peer assessment before reviewing the different methods available.

According to Pond and ul-Haq (1997, p.334) peer review "allows students to influence the distribution to group members of tutor generated marks rather than the generation of marks by students themselves" and focuses on the learning process rather than on the product.

Peer Review is a methodology which allows the student to provide a limited and controlled input into the assessment procedure through

evaluation of each other's performance in the out-of-class group learning activities with control of the final assessment grade awarded remaining with the lecturer.

This mechanism is also referred to in the literature as 'peer assessment,' 'peer evaluation' or 'peer appraisal'. For the purposes of this paper the mechanism allowing translation of performance into a mark or grade in assessed group work will be referred to as summative peer assessment.

Review of peer assessment methods

There are a number of ways in which the contribution of individuals towards assessed group work can be calculated. The group could openly discuss each member's contribution or members could complete an assessment form in secret. This element of peer assessment might form a fixed fraction of each student's overall mark with the remaining fraction coming from the group mark for the product (for example a group project or presentation), or peer assessment might be used to calculate a weighting factor which is used to arrive at individual marks from the group product mark. Assessment might be based on a list of categories (for example Goldfinch & Raeside, 1990; Habeshaw, 1993; Johnson, 1993; Gatfield, 1999; Li, 2001) or be based on a holistic approach involving a single overall score (for example Lejk & Wyvill, 2001a).

Sharp (2006, p.330) states that methods range from the 'minimalist' (i.e. students are awarded the same mark unless intervention is required by the tutor because one or more students are not pulling their weight) through 'abdicatory' (multiplying the group mark by the number of students and leaving each group to decide who gets what) to the 'statistical' (designing some way of combining student evaluations of each others' contributions with the group mark to arrive at individual marks). The following review provides an explanation and critique of the three

most common methods together with an analysis of students' perceptions if reported.

Sharing a pool of marks between group members

This method is described by Gibbs (1986); Habeshaw et al (1993); Lejk & Wyvill (1996); Healey (1999) and Falchikov (2005). Students decide amongst themselves through discussion how to split up a group mark awarded by the tutor. For example if the group mark awarded were 60% and the group consisted of three students (students A, B and C), the group would be awarded a total of 180 marks to distribute amongst themselves (60 x 3 = 180). Students might, for example, decide marks as follows: A: 65%; B: 55% and C: 60%. In some cases criteria for assessment are agreed beforehand and sometimes not.

Strachan & Wilcox (1996) describe this as a zero-sum approach, which:

> allows that in any one group a student's mark may go up or down, as long as the sum of the movements for the group is zero. Thus, no net gain or loss is made from the assigned group mark.

This method is predicated on the notion that if somebody contributed more, then someone else will have contributed less than the given mark deserves.

Strachan and Wilcox's (1996) students were divided into groups of three with each group member contributing to a seminar and essay. Students assessed each other in six areas including leadership and intellectual contribution using a scale from –2 (little or no contribution) to +2 (outstanding contribution). Space was provided for students to justify their marks. They found that some students preferred evaluative comments to number ratings, others did not like "ratting on their friends" and:

> a particular concern was that they were unable to give all group members outstanding or above average ratings with a zero-sum system, even though such ratings would have reflected the actual contribution made by the members of highly effective groups (Strachan and Wilcox, 1996 p.347).

Positive aspects reported by their students included the opinion that the process "weeds out lazy butts", "the process of choosing evaluative criteria as a group was very helpful", "peer assessment helps the instructor make a more accurate assessment of student performance" and "knowing that we would be evaluating each other may act as an incentive" (p.347). They also reported that students in their study appreciated being included in the process of setting criteria and "debated enthusiastically the merits of various criteria" (p.346). However, of the 30 students involved only 15 completed an evaluation form so the sample was small. They point out the need to develop some way to pass on the evaluative comments of peers without divulging the identity of the students involved.

Earl (1986, p.65) rejected this method on the grounds that it breeds competition rather than collaboration as do Conway et al (1993) who argue that it introduces an element of competition into an otherwise collaborative venture. Goldfinch & Raeside (1990, p.211) criticise it for three reasons: firstly, if discussion becomes heated, the resultant bad feelings can have a detrimental effect on students' relationships and subsequent attitude to group work; secondly, if students are reluctant to be critical of their peers, marking can fail to differentiate between students' actual contributions; and thirdly, marks may reflect students' personalities rather than their contributions.

Healey (1999) found that the individually weighted group marks technique was more effective at distinguishing the contribution individuals made to a group project than the pool of marks technique (discussed above) because it resulted in a wider range of marks. This is likely to be due to the fact that students were reluctant to mark their peers down openly but

were prepared to do so if the marks they allocated their peers remained confidential.

Rafiq & Fullerton (1996, p.70) also highlight the problems they faced using this method:

> Early approaches involved the group members negotiating their share of the mark from a total group mark awarded by the lecturer. Some groups would agree to share equal marks, but this was hard on the hard workers, unduly generous to the ineffective and indolent, while the inspired and industrious were not identified. Although it had the dubious value of replicating real life where reward and reputation are dependent on joint effort it was, for many students, a bitter pill. An alternative approach in which students negotiated individual marks amongst themselves fared no better and on occasion led to bitter enmity and shouts of foul-play. Clearly a more equitable strategy had to be found and one in which the value judgements of peers could be made without creating resentment.

It can be seen from this discussion of the literature that students seem reluctant to engage in this method of peer assessment if it is carried out openly.

Weighting the group mark according to individual contribution

This method usually involves the multiplication of a group mark by an individual weighting factor. It was first reported by Goldfinch & Raeside (1990) and then modified by others including Conway & Kember (1993), Goldfinch (1994), Rafiq and Fullerton (1996), Li (2001), Cheng & Warren (2000) and Sharp (2006). It involves the tutor allocating a group mark, which is then manipulated to derive a mark for each individual within the group by taking peers' assessments into account.

One example involves a two-part process involving two questionnaires. Part 1 relates to the skills involved in the group work assigned. It lists all the tasks that the group should have completed during the project so that

the contribution of each member of the group can be distinguished and the relevant individuals identified against each task. The number of times an individual is mentioned is compared to the maximum possible number of mentions.

Part 2 of the questionnaire summarises a list of process skills related to group activities such as those below decided on by tutors in the Goldfinch and Raeside (1990) study:

- Level of participation
- Suggesting ideas
- Understanding what was required
- Helping the group to function well as a team
- Organising the group and ensuring things got done
- Performing tasks efficiently

Students are required to award a mark, for example of between 0 and 4 (Goldfinch & Raeside, 1990), or –1 and 3 (Goldfinch, 1994) to each group member to reflect the proportion of the student's contribution to that part of the process. The score for Part 2 is the result of the actual number of marks allocated to an individual, divided by the total possible marks.

Conway and Kember (1993) explain how marks are calculated (see Table 5.1):

Part 1 score = number of mentions of individual / possible number of mentions Part 2 score = actual sum scored / highest possible score Part 1 and 2 scores are then combined by the formula: Peer Assessment (PA) score = Part 1 score x 1/3 + Part 2 score x 2/3 The PA score is then converted to a PA factor (a %) using a table. The PA factor is then used to weight the awarded group mark.

Table 5.1 Conway and Kember's (1993) method of calculating a student's individual mark using a weighting method.

The process is usually administered under exam conditions; non-attendance can result in a deduction of marks and marks allocated are kept confidential. Goldfinch and Raeside found there was an "improvement in class spirit" (1990, p.222) and that there was considerable agreement between members of a group about peers on both Part 1 and 2. They also found that the final ranking of peer assessment scores within a group matched lecturers' expectations (1990, p.218) although, since lecturers did not attend group meetings, it is difficult to see how this could be possible. Students' perceptions are not discussed.

Conway and Kember (1993) found three main drawbacks to this method in their research. Instead of giving the names of individual students who had made a major contribution, students cited "everybody" in 67.9% of cases. Secondly, the 0 – 4 scale was not successful; nobody awarded a 0 which meant the mean was 3.0, well above the stated "2 = average" (p.48). Thirdly, the calculations were lengthy and tedious.

There have been several other versions similar to this method reported in the literature since Goldfinch and Raeside introduced it in 1990. For example, Habeshaw (1993) suggested using the matrix shown in Table 5.2.

	Major contribution	Some contribution	Little contribution
Leadership and direction	0	-1	-2
Organisation and management	0	-1	-2
Ideas and suggestions	0	-1	-2
Data collection	0	-2	-4
Data analysis	0	-2	-4
Report writing	0	-3	-6

Table 5.2 Goldfinch and Raeside's (1990) peer assessment matrix.

A group member who makes a major contribution in all areas scores a rating of 0 and gets the group mark but a peer who makes little effort in all areas would get 20 marks deducted from the group mark. The benefits of this method are, firstly, that different weightings can be attached to different aspects; and secondly, no student gets more than the group mark. However, one drawback could be that, since students tend to dislike giving negative marks, they record all members as having made a major contribution. Gibbs (1992) suggests a variation with marks from –2, –1, 0, +1, +2 so that it is possible for members who have contributed well to get more than the group mark as long as the average of the given marks equates to the group mark. This method, however, could create at least three problems: firstly, the use of negative marks is still evident; secondly, to reward some group members students would have to penalise others which could prove unpopular; and, thirdly, it does not overcome the problem highlighted by Conway and Kember below (1993, p.46/47) that students could get a high mark for producing a poor piece of work.

Other versions of this process include the use of a Likert Scale for example the student-generated instrument reported by Johnson (1993). In this case students identified thirty characteristics of working as a member of a group which were honed down to 13 using a quasi-factor analysis (for example attendance at group meetings; meeting group deadlines). Students were then asked to complete a form stating to what extent they agreed each other member of the group met each of the criteria using the scale indicated in Table 5.3 below:

Strongly agree	Somewhat agree	Not sure	Somewhat disagree	Strongly disagree
5	4	3	2	1

Table 5.3 Johnson's (1993) Likert scale used for peer assessment of group work skills

Each student's mark was then calculated as the average of the total marks awarded by other group members.

Gatfield (1999) found high levels of satisfaction with a weighted assessment method he adopted. This involved a tutor mark for a group project (worth 50%) and peer assessment marks for the process (50%). Criteria for peer assessment were agreed with a previous cohort. Students agreed that students should assess their peers; it was a fair method to arrive at marks; the system reflected students' efforts; and, peers can assess fairly.

Although Cheng & Warren's (1997) work elicited student perceptions before and after a peer assessment exercise using this method, questions did not differentiate between what students thought of assessment of product (presentations and a written report) and assessment of contribution. They found that about 66% of students believed that students should participate in peer assessment and that although the majority of students had a positive attitude towards it, they were also unsure or negative towards the idea of first year students actually participating in it. Before the peer assessment exercise the majority said they would feel uncomfortable undertaking peer assessment but afterwards almost half said they did feel comfortable. Cheng & Warren suggest that more opportunities for peer assessment could reduce the number who feel uncomfortable concluding: "the implementation of peer assessment alone goes some way towards dispelling students' initial reservations" (1997, p.236). These findings are in line with findings by Burnett & Cavaye (1980), Earl (1986), and Williams (1992).

In Pond & ul-Haq's (1997) study, students were divided into groups which presented written and verbal solutions to cases in banking law and practice each week. Tutors marked the product while the group being assessed completed a peer assessment form each week to weight the marks awarded by the tutor amongst group members. Students' assessed their peers on a scale of 0 (no contribution) to 4 (very significant contribution) in terms of individual participation, team spirit, research/ preparation and contribution to tutorials. Discussion of the

students' perceptions of the process is limited but Pond & ul-Haq do report that students felt peer assessment of group work was an effective tool and that its use should be continued.

A major problem can be identified with this method in terms of the weighting of marks. As Conway & Kember (1993) state:

> care needs to be taken over the relative weighting of the base and effort marks. If the effort marks are high compared to the base marks, students who make a greater contribution than their fellows to a project which is skimpy or fatally flawed are likely to end up with a better mark than a lesser contributor to an outstanding project. If contribution marks are weighted highly, a subtraction procedure is likely to fail students unless they produce very good projects. On the other hand, low weighting to the effort mark could reduce its significance to the extent that students will complain that they were not rewarded for their effort (Conway & Kember, 1993, p.46/7).

Although Conway & Kember's (1993) method went some way towards simplifying Goldfinch and Raeside's (1990) original method (1990), and Li (2001, p.7) introduced a normalisation process to "iron out inherent shortcomings," according to Sharp "there is still some dispute about what ratings should be collected and how they should be collected" (2006, p.330). Sharp (2006) suggests that if adjustments are to be made based on individual contribution to a group mark, they should be made on sound statistical grounds.

Separation of process and product

Falchikov, (1988 and 1991) and Bloxham & Boyd (2007, p.98) distinguish between the assessment of the product (carried out by the tutor) and the process (carried out by the students' peer/self assessment). Instead of one mark arrived at using a weighting factor as described above, each student receives two marks, one for the product and one for the peer assessment of contribution to the process. Falchikov (1993) suggests detailed criteria for assessment, divided into task functions and group

maintenance functions (developed from the work of Johnson & Johnson, 1975) shown in Table 5.4 below:

Task Functions	Group Maintenance Functions
Information and Opinion giver	Encourager of Participation
Information and opinion seeker	Harmoniser and Compromiser
Starter	Tension Reliever
Direction giver	Communication Helper
Coordinator	Process Observer
Diagnoser	Standard Setter
Feasibility Tester	Active Listener
Evaluator	Trust Builder
(Link person)	
(Time keeper)	

Table 5.4 Falchikov's (1993) criteria for assessing group work skills

Falchikov added two open categories which students could decide for themselves. Students assessed each other as high, medium, or low in all these areas and marks are awarded accordingly (although as Lejk & Wyvill (1996) point out, she does not make it clear how this is done).

Johnson (1993) uses a similar system except that the criteria are generated by the students themselves. In both cases the students receive two marks – a mark for the product (the same for each group member) and a peer assessment mark reflecting their contribution. The latter is arrived at by averaging the marks awarded by other group members as with the weighting method described above.

Krause and Popovich's (1996) research involved separation of process and product. They developed, implemented and evaluated an anonymous peer assessment (of contribution to group work) system with a fourth year undergraduate pharmacy course involving 152 students (divided into 28 groups with 5 or 6 in each group). Students were asked to rate peers on ten statements (for example: this person actively contributes to group discussions, e.g. provides ideas, shares insights) on a five point scale.

Students were encouraged to include constructive comments on the feedback sheets on each person and this was then typed up by the tutor to ensure anonymity. The process was carried out twice, eight weeks into the course and at the end. On both occasions each student was provided with a feedback sheet which included the average mark awarded by his/her peers based on assessment of the ten areas which they could then compare with their self-assessment. Any qualitative feedback provided by other group members was also included. Krause and Popovich found that 71% of students involved in the research believed that peer assessment of contribution increased their accountability to other group members and 80% believed receiving constructive feedback from peers helped them identify strengths and weaknesses of their group work abilities. They claim that in order to encourage students to be objective when assessing their peers, tutors must ensure that the feedback is presented to students "in an anonymous, constructive manner"(Krause and Popovich, 1996 p.142).

Conclusion

Section 1 of this chapter reviewed the literature on social loafing and found 'identifiability' to be a key factor in alleviating social loafing. Section 2 explored a number of ways in which the contribution of individuals towards assessed group work can be calculated. If the decision were made to introduce a peer assessment mechanism to arrive at individual marks for students' contribution to group work, one of the above methods would need to be adopted. Before making a decision, however, it made sense to make sure that students felt strongly enough about social loafing to introduce summative peer assessment to try to alleviate it. Although research with the 147 first year students reported in Chapter 4 suggested that students did, it was important to extend my sample to include second and third years students in order to warrant my assertions for action (Morgan, 2007).

It was also apparent from the review of the literature in Section 2 that there were two gaps in research reported in this area. Firstly, although there are a number of studies which describe methods of using peer assessment of contribution to group effort to arrive at an individual mark, there is very little discussion about what students think of such methods (Hanrahan & Isaacs, 2001; Struyven et al, 2002). Secondly, there is little discussion about how any qualitative, formative feedback is communicated to the students except for the work of Krause & Popovich (1996). If the pool of marks method were to be used, how effective would the discussion of marks be in terms of formative feedback? Although there is an obvious need for a summative mark, students need clear feedback so they can identify the strengths and weaknesses of both their group working skills and the product produced.

Reports of peer assessment which include student perceptions tend to concentrate on the process of peer assessment of product (for example, Cheng & Warren, 1995) and focus on the researcher's point of view (for example, Mello, 1993; Healey, 1999), the mechanics of such systems such as how to calculate weighting factors or establish rank order (e.g. Li, 2001; Bushell, 2006; Sharp, 2006, Kilic & Cakan, 2006), the validity of peer assessment (e.g. Falchikov, 1986; Kwan & Leung, 1996; Cheng & Warren, 1999 & 2000; Magin, 2001) or the merits of using holistic versus category-based approaches (e.g. Lejk & Wyvill 2001a, 2002). The limited discussion that exists about students' perceptions of using peer assessment to identify individuals' contribution to group work is usually quantitative (for example Gatfield, 1999; Pond & ul-Haq, 1997) or assigned a few sentences in a report dedicated to a study of the process involved from the researcher's viewpoint (for example Butcher & Stefani; 1995).

This encouraged me to start on the second cycle of action research to find out whether students would choose to implement summative peer assessment; if they did, which of the methods discussed above they

would prefer; how formative feedback could be incorporated; and to evaluate their perceptions of both summative and formative processes.

Chapter 6: Report of findings about student choices of methods of assessing group work

This chapter discusses the findings from two samples of students in terms of their choices of methods of assessing group work. It is divided into three sections. Section 1 briefly outlines the method used to ascertain whether students would prefer to receive one product mark for assessed group work or for the group mark to be individualised by assessing each other's contribution to the group task. It then discusses the findings. Section 2 reports the research undertaken with a second sample of students to ascertain which of three options they would prefer: peer assessment of contribution to group effort; distribution of a pool of marks, or; equally shared product mark with exceptional tutor intervention. It then focuses on the choices made by social loafers and discusses possible reasons for their choices. Section 3 synthesises the findings, draws conclusions and suggests questions that need to be resolved before the third cycle of the action research process is undertaken.

Section 1: Sample 1 - Method and discussion of findings

Method

Thirty four second and third year undergraduates working in groups of three or four on an education module were asked to decide which of two assessment options would be more effective in terms of ensuring individual accountability: a group mark weighted according to summative peer assessment of contribution to group effort or one group mark for product. The concept of individual accountability in group work had been discussed in some depth with students beforehand. Students were asked to give reasons for their answers using a short questionnaire (Questionnaire 2). The aim was exploratory; a first (pilot) attempt to

ascertain students' general feelings about the subject of individual accountability.

Discussion of findings

Thirty-one students completed the questionnaire (response rate 91%). Thirty students opted for the first option: summative peer assessment. This is illustrated in Figure 6:1.

Figure 6.1 Number of students opting for peer assessment of contribution to group effort (n = 31)

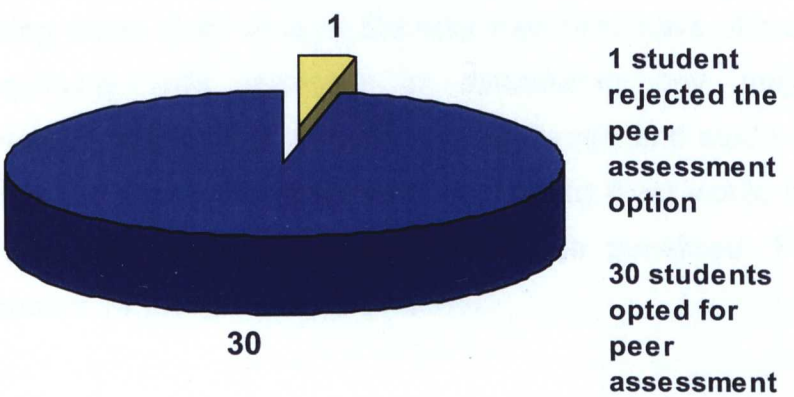


Figure 6.1 Pie chart indicating the number of students choosing peer assessment of contribution rather than sharing one group mark for product.

Sixteen of the 31 respondents (52%) felt that summative peer assessment of contribution would make students work harder. For example Respondent 1 stated:

> This would be very effective as the members of the group would want to be seen to be avoiding social loafing and would want to be seen as a productive member of the team.

Respondent 6 had already experienced summative peer assessment in a first year module and:

> found the idea that my group would be assessing me affected how I worked. I believe that this is a good way of stopping social loafing as not only do you want your group members to grade you highly so you do well in the module, but you want your group to have recognised your efforts and appreciate your work.

The second reason students chose the summative peer assessment option was to do with a notion of validity. The twelve students (39%) commenting on this aspect either acknowledged the fact that the students were best placed to assess contribution, for example, Respondent 12, who stated: "I think it's very effective as it is the group members who know how hard each other have worked towards the project," or liked the way that multiple viewpoints increased accuracy. For example, Respondent 18 thought it was "good because the marker can see correlations between member responses." Respondent 21 felt it was "effective in getting more of an idea of the way members have shared workload .....certainly aids assessor in discovering any major anomalies." Increased validity was also seen to be "fairer" and students also liked the idea that group members who contributed most would be rewarded and those who contributed least would be penalised. For example, Respondent 14 thought the process was:

> fairer as the contribution of individuals is accounted for and so anyone who has provided less input can be penalised appropriately and other members rewarded for their individual efforts.

Respondent 25 confirmed: "It makes sure the marks are fair and that everyone's effort is taken into consideration."

However, despite the fact that 30 out of 31 students felt that peer assessment of contribution to group effort was a more effective way of structuring individual accountability, there were a lot of 'ifs and buts'. Fifteen out of the 31 respondents (48%) were concerned about biased marking by peers. For example, Respondent 16 expressed concern that:

> if the students are friends/have formed some form of relationship they will not want the others to fail and may mark them higher than the amount of work they have completed will merit.

Respondent 21 was concerned that "more popular members may be scored higher regardless of work" while Respondent 25 was worried what would happen when "group members don't get on."

Despite their reservations, these students still felt that peer assessment of contribution was a better option than being awarded one product mark and only one of the 31 students said that awarding group members all the same product mark was a better option.

Although most students did not consider whether peer assessment would be anonymous or not, three students did broach the issue. Respondent 4 thought peer assessment was "good because you can say if you feel others haven't put as much effort in without directly confronting them" so obviously assumed the process would be anonymous. Respondent 5 stated: I think that the individual assessment of contribution should be anonymous as if conducted together it could cause unwanted conflict" and Respondent 10 agreed with peer assessment "only as long as the other group members cannot see your individual assessment of your peers."

In summary, thirty out of the thirty-one respondents stated that peer assessment of contribution to group effort was a more effective way to structure individual accountability. They felt that it would make students work harder and would increase validity of assessment since the tutor was not in a position to assess individuals' contributions. They thought this was fairer although 48% had concerns about how biased the marks would be. Four students suggested that if marks were confidential, students were more likely to be honest about their peers' contributions.

However, asking students studying assessment which mechanism *would be* most effective in structuring individual accountability is not the same

as asking them which method they *would actually choose*. Hypothetical questions lead to answers that lack validity and students are more likely to give serious consideration to questions that relate to their own welfare. So the next stage of the research involved students whose choices had an impact on their own assessment.

Section 2: Sample 2 - Method and discussion of findings

This section briefly describes the method used before reporting the findings of research with a second sample of students. The following research questions are addressed:

- Which of three choices (peer assessment, distribution of a pool of marks or equally shared mark with exceptional tutor intervention) would students make and why?
- Is there any significant difference in the choices made by social loafers and those not involved in social loafing?
- What might account for these differences?

Method

Seventy-four second and third year undergraduate education students working mostly in groups of four on a group assignment were given the opportunity to choose how they would be assessed. The three options in Table 6.1 were introduced and discussed. Students were asked to give written reasons for their choices on a short questionnaire (Questionnaire 3).

1. **Peer assessment of contribution to group effort.** The tutor gives a mark for the group project, which is then modified according to peer assessment of contribution to the group effort. Peer assessment marks are awarded based on criteria agreed on by the group at the beginning of the module. The marks awarded by peers are averaged to arrive at a mark and remain confidential.

2. **Distribution of a pool of marks.** Students receive a group mark (for example, 60%) for the project from the tutor. The group mark is then multiplied by the number of members in the group (for example, three members in a group= 60 x 3 = 180 marks.) Students then openly discuss and agree how to divide the 180 marks between themselves, for example: student A might get 70%, student B 60% and student C 50%.

3. **Equally shared mark with exceptional tutor intervention.** Each member of the group is awarded the same group mark unless the tutor is approached because one or more students are felt not to be contributing. A meeting between the tutor and all members of the group takes place should this happen, the problems are discussed and individual marks agreed between the group and tutor.

Table 6.1 Choice of assessment methods offered to students

Discussion of findings

Sixty-six students returned questionnaires (89% response rate). Table 6.2 shows the choices made by students.

Choices available	Number of students
Method 1: peer assessment	42 (63.6%)
Method 2: distribution of pool of marks	5 (7.6%)
Method 3: group mark	19 (28.8%)
	Total: 66 students

Table 6.2 Students' choices of assessment methods

Reasons for choosing Method 1 (peer assessment)

Eight themes emerged from the data in response to Method 1 (peer assessment of contribution to group effort), which related primarily to respondents':

- belief that peer assessment of contribution may discourage social loafing (or identify social loafers);
- acknowledgment of the importance of differentiating between the product and process of group work;
- belief that the tutor is not in a position to assess the process of group work; group members are better placed to assess peers' contributions;
- belief that peer assessment of contribution is "fair".

They did, however, also acknowledge:

- that marks awarded to peers could be biased due to friendship/dislike.

A number also expressed the feeling that:

- confidentiality when awarding marks to peers would ensure honesty;
- tutors should assess the product.

They also expressed:

- a feeling of ownership of the process.

Belief that peer assessment of contribution may discourage social loafing (or identify social loafers)

Seventeen students (40%) identified the capacity of peer assessment to prevent social loafing or identify those who did not pull their weight. Students felt that the capacity of peer assessment to make peers feel "identifiable" or "accountable" was important. For example, Respondent 16 felt "individual contributions are being *identified* by peer assessment and therefore students feel their work is being recognised, preventing social loafing," while Respondent 16 stated: "Peer assessment also *identifies* those who let the group down" and Respondent 45 suggested that "it allows each individual to feel *accountable* for the work they produce hopefully reducing the effect of social loafing." Students also pointed out the importance of Harkins and Jackson's (1985) evaluation potential, for example Respondent 31 suggested peer assessment "highlights members who have worked particularly well and ones who haven't done as well."

Acknowledgment of the importance of differentiating between the product and process of group work

Fourteen students out of 42 (33%) acknowledged the need for differentiation between the process and product of group work for example Respondent 1:

> I like this method because each individual's mark can vary depending on their effort and contribution as assessed by peers but it is also based on the group mark for the overall work produced.

Gillies and Ashman 2003 (p.80) underline the importance of assessing the entirety of the educational experience so that students do not ignore the learning process and focus only on the learning outcome. This holistic approach is reflected in the comment of, for example, Respondent 46:

> Since the success of group working is determined not only by the quality of the final product but also by how effectively the individuals function within the group, it seems reasonable to include within the assessment an element that awards marks for group maintenance functions and task functions rather than solely for the final product.

Belief that the tutor is not in a position to assess the process of group work; group members are better placed to assess peers' contributions

Twelve out of 42 students (29%) acknowledged that group members were in a better position to assess the process of group work, for example Respondent 9 asserted: "Only the group members can adequately acknowledge who has put in the most and least work and it would be wrong to ignore this essential fact" while Respondent 22: reported:

> I think that when working in a group the opinions of the group members can be just as, if not more, important than that of the teacher in terms of judgements. Students who appear conscientious and hard working in class may work poorly and contribute little in a group situation.

Belief that peer assessment of contribution is "fair"

Ten students (24%) believed peer assessment to be "fair" for example Respondent 4 regarded it as: "*fair* because social loafers will not have the same mark as the ones who have contributed more to the group work" as opposed to it being "*unfair* for a tutor to give an overall mark to a group without the input of the group members" (Respondent 29). Respondent 39 felt it was "a very balanced and *fair* method of assessing work. Students have their say about each other's participation levels but are still encouraged to work together for a common aim."

Acknowledgement that marks awarded to peers could be biased due to friendship/dislike

However, twelve students (29%) expressed concern that marks awarded to peers could be biased by friendship or dislike, for example, Respondent 2 conceded: "More often than not people dislike giving their friends bad grades no matter how much work they have or have not put in."

Some students acknowledged the possibility of biased marking by peers and still opted for Method 1 perhaps because, as Respondent 4 concluded, they could not "see a better solution to assessing group mates than the solution in Method 1."

Confidentiality when awarding marks to peers would ensure honesty

Three students (7%) expressed the feeling that confidentiality when allocating marks would ensure honesty. For example Respondent 3 claimed: "If the peer assessment is confidential then I think all group members will be honest even if they are friends; it is more likely if people are unhappy with a group member that they will speak out about it."

Tutors should assess the product

Six students (14%) reflected the belief that although peer assessment was an effective means of assessing the process, the tutor should be responsible for assessing the product, "because they understand the work on a much deeper level" (Respondent 32) being "independent and expert" (Respondent 43), tutors "are trained to assess students' work" (Respondent 58) and "know what they are looking for in a piece of work" (Respondent 35).

A feeling of ownership of the process

Six students (14%) expressed a feeling of "ownership of the learning we are undertaking" (Respondent 58).

Reasons for choosing Method 2 (distribution of a pool of marks)

Five of the 66 students (8%) chose Method 2 (distribution of a pool of marks). Respondent 18 felt that Method 2 would mean "each individual would put in much effort so as to gain a high group grade then divide it among themselves as to who did most of the work."

Respondent 19 believed Method 2 to be:

> the fairest method of all, with a set amount of marks to begin with each group member can then distribute them to their peers as they see fit. This method will also force each group member to think very carefully about where they put their marks because each mark will have an effect on the marks they give to another group member. My only slight reservation about this process is that it may be slightly restrictive and therefore result in group members being more conservative in their grades than they other wise might have been with another method for instance Method 1 would probably result in a wider spread of grades.

Respondent 21 chose Method 2 reluctantly but felt none of the methods:

> takes into account individual effort without being chosen by the group which I feel would cause conflict or biased results as friends within a group may not feel comfortable with giving peer feedback in an honest way. Method 2 would be the best – it would be a fair reflection of how much each student had contributed, however this could lead to arguments when agreeing on what percentage each student should receive.

Four students who did not choose Method 2 criticise it for precisely this reason (difficulty over agreeing marks). For example:

> If the group were to distribute marks, this could cause tension and arguments within the group as individuals might feel hurt and upset if

some members said they haven't done enough work, or if people feel they should have got a better mark than the group are suggesting. (Respondent 5)

> ....each member might disagree with the way they would like the marks distributed. (Respondent 13)
>
> In the distribution of a pool of marks, competition as well as mortification of weak students may be facilitated as everybody's individual marks deviates from the group mark in a very direct i.e. positive or negative way and individual deficits become very clear. (Respondent 17)
>
> If it is up to the group to distribute marks, there may be a great deal of bias and conflict involved in this process. (Respondent 57)

Reasons for choosing Method 3 (all students receiving the same group mark)

Nineteen students chose Method 3 (29%) - equally shared mark with exceptional tutor intervention. The main reason students gave for choosing this method was to avoid the possibility of peers' biased marking. For example:

> I do find the separation of process and product (Method 1) a good idea however I went against it as I feel that if you are all good friends in a group it can be hard to mark their work and effort completely honestly." (Respondent 53)

The other main reason for choosing this method was because group members had worked together before or knew and trusted each other, for example:

> Having worked with the other group members previously I know that we are of equal standing in regards to both effort and capability. (Respondent 61)
>
> It turned out we all kept to deadlines and agreements. Also we are all very committed and have productive discussions concerning the group assignment. I know I can rely on the others to get the work done. (Respondent 65)

What method did social loafers choose?

In this pilot research, a peer assessment mechanism was introduced and social loafers were identified. They were categorised as those students receiving a consistently low score from their peers for contribution to group effort. It was possible to identify social loafers because although students were given the opportunity to remain anonymous, most elected to include their names on completed questionnaires.

The process of identifying social loafers was not undertaken until after qualitative analysis had taken place. The reason for this was to maintain objectivity by avoiding any subconscious attempt to bias how responses were categorised.

Of the 66 students who completed questionnaires, 11 students (17%) were classified as social loafers. In addition, 6 of the 8 students who did not attend the final session and therefore did not complete questionnaires were also identified as social loafers (perhaps lending more weight to their identification as non-contributors). Table 6.3 shows the choices social loafers made.

Choices available	Social loafers' choices
Method 1: peer assessment	1 out of 42
Method 2: distribution of pool of marks	1 out of 5
Method 3: group mark	9 out of 19
	Total: 11 out of 66 students

Table 6.3 Choices of assessment method made by social loafers

A contingency table was formulated showing observed and expected frequencies for choices of assessment method. Results showed that $p < 0.05$. Two results are striking here:

1. Significantly fewer social loafers chose Method 1 (peer assessment) than can be accounted for by chance.

2. Significantly more social loafers chose Method 3 (equally shared mark with exceptional tutor intervention) than can be accounted for by chance.

Of the eleven social loafers who completed questionnaires, nine chose Method 3 (group mark), which is the method least likely to identify a social loafer. Analysis of their responses showed that either the issues identified by those not classified as social loafers were not discussed (or perhaps avoided) or were counter to those expressed by students not identified as social loafers. For example Respondent 56 claimed that "the tutor should also be more qualified at marking a group's efforts," despite the fact that tutors are not present during group meetings, a fact acknowledged by a number of students.

A number of students *not* classified as social loafers asserted that individuals rarely exert equal effort in group situations (this is mentioned 38 times). Despite this, all nine social loafers expressed the belief that it is possible for work to be shared equally or that other group members can be trusted to play an equal role (while failing to do so themselves, according to their peers). For example Respondent 48 claimed: "the vast majority of the time students all contribute equally to group work therefore all deserve the same grade." Respondent 49 states: "I feel that we should be organised enough that we all ensure we have put in equal amounts of effort/work towards this assignment." Respondent 51 opts for a shared grade "because each member had to contribute to the work to achieve the end result." An alternative explanation is that these students realise that whereas their fellow group members are unlikely to "shop" them directly for failing to pull their weight, they might be likely to do so using a confidential summative peer assessment form. In other words, Method 3 (a shared grade) is their best option in order to avoid being penalised for social loafing.

Section 3: Conclusion

This chapter reported the findings from two groups of second and third year students (n = 34 and 74 respectively). Research with the first group explored whether students would prefer to receive one group mark for assessed group work or for the group mark to be individualised by assessing each other's contribution to the group task. Thirty out of the 31 respondents opted for the latter. Research with the second sample of students sought to ascertain which of three options they would prefer: peer assessment of contribution to group effort; distribution of a pool of marks or equally shared mark with exceptional tutor intervention. The majority of students (64%) chose the first option.

The main reasons for students' choices in the second sample related primarily to respondents' perceptions that peer assessment of contribution may prevent and will identify social loafing; acknowledgment of the importance of separation of product and process of group work; the belief that the tutor is not in a position to assess the process of group work; and that peer assessment of contribution is "fair".

The reason why Method 2 (distribution of a pool of marks) was unpopular with students was mainly due to the conflict students imagined would occur as a result of having to openly agree marks.

Method 3 (one group mark) was chosen by students who claimed to trust their peers. However 9 of the 19 students who chose this method were identified as social loafers whose trust was not earned. A significant number of social loafers (x^2 (2) = 18.87, $p < 0.05$) chose Method 3. It seems likely that they chose this method because their previous experience of group work led them to believe that their peers would be reluctant to involve a tutor, should problems with a group member occur. Studies by Strachan & Wilcox (1996); Healey (1999); Cheng & Warren (2000); Lejk & Wyvill (2001b); McWhaw et al (2003) and Barfield (2003)

support this claim. These studies found that students tend to put up with a social loafer rather then report him/her to the tutor.

However, although students felt that peer assessment helped alleviate social loafing, there is no proof that this was so. Although social loafers were identified (consequently the system did not prevent their social loafing) there is no way of knowing whether other students were prevented from becoming social loafers because of the intervention.

As a result of this research I decided to implement a summative and formative peer assessment system. Although some students expressed concern over how biased marks would be, the majority of students chose this option. However, before undertaking this intervention there were various questions that needed to be addressed:

- Should peer assessment be based on one holistic figure or a number of pre-specified criteria?
- If the latter, who should decide on the criteria to be used?
- Should self-assessment be included or excluded?
- If the allocation of summative marks were to be confidential, how could a formative element be incorporated so that students receive the benefit of feedback to facilitate improvement of their group working skills?
- How exactly should summative marks be calculated?

The above questions are addressed in the next chapter.

Chapter 7: Questions to be considered before implementing peer assessment

Research with two groups of students reported in the last chapter led to the decision to implement summative peer assessment in order to identify individual contributions to group work. This chapter addresses the questions identified in the previous chapter which needed to be answered before peer assessment could be implemented.

Section 1 explores whether summative peer assessment should be based on one holistic mark or on a number of pre-specified criteria. Section 2 discusses who should decide on the criteria to be used if criterion-referenced assessment were chosen. Section 3 considers whether self-assessment should be included or excluded. Section 4 explores whether or not the allocation of summative marks should be confidential. Section 5 reflects on the importance of formative peer assessment. Section 6 explores how a formative element could be incorporated so that students receive the benefit of feedback to facilitate improvement of their group working skills and discusses whether or not feedback should be anonymous.

Section 1: Holistic versus criterion-based ratings. Should peer assessment be based on one holistic figure or a number of pre-specified criteria?

According to Ritter (1998, p.79), "Effective use of peer assessment ... seems to depend on ensuring [students'] understanding of the criteria being applied and their competence to apply them...." Yorke (2003, p.487) claims that "students' understanding of the assessment task is enhanced through the specification of assessment criteria."

There is much evidence in the literature to support criterion-based assessment. Fry (1990, p.178) for example, found that in the absence of clear criteria students tended to use "idiosyncratic and misconceived criteria." Stefani (1998), Falchikov (1998, p.10) and Klenowski (1995, p.148) emphasise the importance of criteria being clear, explicit, negotiated and understood. Brown et al (1998, p.111) and Topping (1998) suggest that students prefer specific performance criteria couched in accessible language to "fuzzy" (Otter, 1995, p.45) vague or holistic ratings (see also Li, 2001, and Conway & Kember, 1993). Ninety per cent of the respondents in Peters' (1996) study believed that it was important for assessment to take place against clearly stated criteria while the students in a study carried out by Sambell et al (1997, p.364) felt that clarity was a fundamental requirement of a fair and valid assessment system.

However Lejk & Wyvill (1996) advocate the superiority of an holistic approach to assessment. They argue that it is easier to capture a general impression of someone than to evaluate him/her on specific dimensions and come to the conclusion that "group assessment is a reliable instrument .....if the overall impression serves as the main criterion for decision-making" (p.274). However, Lejk and Wyvill's study is based on the work of Schechtman (1991 and 1992) whose research was based on admissions tutors' agreement as a group whether or not to accept students for initial teaching training degrees in a higher education setting. These tutors have what Rust et al (2003) refer to as 'tacit' knowledge of what they are looking for in a candidate. This is very different to students, as novice assessors, trying to evaluate the contribution of other group members.

Adopting an holistic approach to students' assessment of peers' levels of contribution would require a tacit understanding of what an effective group member might be. Rust et al (2003, p.152) argue that tacit knowledge is "highly personal and hard to formalise" and is based on "our ingrained mental models, beliefs and perspectives." Webster et al (2000,

p.73) liken holistic assessment to "wine tasting – a high level activity that requires continued practice but that is pretty much impenetrable to non-cognoscenti." Rust et al (2003) and Yorke (2003) claim that tacit knowledge or "connoisseurship" (Polanyi, 1958, p.54) can only be acquired through discussion, observation and practice over a period of time. This interpretation suggests that holistic assessment is unlikely to be effective when used by students because it does not facilitate the gradual process of internalisation of standards. If students are to understand assessment, they need to be socialised into such a process and the use of specific criteria makes this more likely.

It would seem also that the use of holistic assessment would prevent discrimination of performance (Miller 2003). Miller's study describes an approach in which assessment was altered from looking at a few, generic aspects of performance to multiple, very discrete aspects of performance. He found that increasing the number of criteria decreased the mean scores and increased the standard deviations of peer and self-assessment marks and that correlation between peer and self-assessment was improved when more specific criteria were used.

Sharp (2006, p.331) confirms Miller's findings by suggesting:

> While a student may have a very high or low opinion of another student's rating on one criterion, it is less likely that that the student will maintain such a high or low opinion across a number of categories. By the same token however, while a student may award an average weighting for a single holistic judgement, it is less likely that he or she would award exactly the same average rating over a number of categories.

Adopting an holistic approach, then, would seem to be a less effective choice but I was interested to see to what extent my students would agree with the choice of a criterion-referenced approach. This is reported in Chapter 8.

Section 2: Should a criterion-referenced system be chosen, who should decide on the criteria to be used?

In 1988, Heron (pp.79 –90) wrote:

> Unilateral control and assessment of students by staff means that the process of education is at odds with the objectives of that process. I believe the objective of the process is the emergence of an educated person: that is a person who is self-determining – who can set his [sic] own learning objectives, devise a rational programme to attain them, set criteria for excellence by which to assess the work he produces, and assess his own work in the light of these criteria ..... But the traditional educational process does not prepare the student to acquire any of these self-determining competencies. In each respect, the staff do it *for* or *to* the students. An educational process that is so determined by others cannot seriously intend to have as its outcome a person who is truly self-determining. (my italics)

Heron's comments draw attention to the fact that students are often denied ownership of assessment, tending to see themselves outside the assessment process and view assessment as something that is done to, on or for them rather than with them. Ecclestone & Swann (1999, p.382), for example, found that students in their study had difficulty conceiving they had an integral role to play in assessment. They argue that students need to be involved in the assessment process if they are to become part of an assessment community.

There are a number of writers who criticise externally imposed assessment criteria and underline the importance of sharing the responsibility of assessment with learners (for example, Claxton, 1995; Peters, 1996). Pain et al (1996) suggest that negotiating assessment increases students' acceptance of it and promotes student reflection. Stefani (1998, p.340) argues that since assessment defines the attitudes students take towards their work, "it is not unreasonable to suggest that all assessment of learning should be carried out in partnership." Claxton (1995), Searby & Ewers (1997), Dochy & McDowell (1997), Hall (1995, p.2) and Boud (1989, p.25) claim that it is of key importance that students

not only carry out peer and self-assessment but are also responsible for determining the criteria on which the assessment is based. Ross and Rolheiser (2003, p.130) and Bryan (2004, p.117) stress the importance of involving students in the setting of criteria for assessment because such practice contributes to a shared language leading to shared understanding of assessment criteria. This in turn increases the likelihood of students acquiring the 'tacit' understanding referred to by Rust et al (2003) and being socialised into the assessment community proposed by Ecclestone and Swann (1999) and Price et al (2008).

There are a number of academics who refer to the empowerment brought about by involving students in the assessment process using words such as 'ownership', 'power' and 'control.' The students in Sambell et al's research (1997, p.364) for example could see the benefits of being involved in the assessment process. "From the student viewpoint, the clarity and openness of such assessment was perceived as an issue of *control*, affording them a measure of independence.... " (My italics.)

Searby & Ewers (1997), Pond and ul-Haq (1997), Falchikov (1998) and Sivan (2000) report that students' ownership of the criteria used for assessment is of fundamental importance:

> Involvement of students in setting the criteria was found to be an essential strategy to maximise the potential of peer assessment for developing students' sense of *ownership and control* .....and for allowing them to exercise responsibility for their learning (Sivan, 2000, p.202). (My italics.)

Cowan (1998, p.87) also found that being involved with the formulation and setting of criteria gave students *ownership* and intimate understanding of their goals which resulted in learning being purposefully directed toward them.

Klenowski (1995, p.146) argues that the ability to assess is integral to the learning process claiming: "students' commitment to learning is likely to

be strengthened when they take more responsibility." Klenowski (1995, p.148) claims:

> When the teacher, together with the students, identifies or discusses the criteria for evaluation, there exists the opportunity to raise the standards of achievement through the clarification of expectations and the explicit statement of performance outcomes to be achieved.

Falchikov (1998, p.19) claims that many practitioners now share the belief that higher education should:

> actively encourage the development of students' abilities to evaluate their own work and that of peers. Such abilities can be cultivated only by involving students in assessment.

She also argues that reliability of marks increases with the number of markers involved whether staff or students. This is supported by Langan et al (2005, p.30) who noted that students who had attended a meeting to develop assessment criteria "marked more accurately."

My intuitive belief that students should be involved in the creation of assessment criteria arises from my social constructive epistemology. Tacit knowledge of assessment criteria and standards is acquired through socialisation into the assessment process. The traditional model whereby tacit standards are absorbed over time informally and serendipitously (Price et al, 2008) is no longer relevant due to the increase in student numbers and the lack of one-on-one tutoring. I made the decision to give the students the opportunity to discuss and design their own criteria in the next cycle of the research should they decide to use criterion-referencing rather than holistic assessment to evaluate their peers contribution to group work.

Section 3: Should each student's self-assessment be included or excluded?

The issue of whether or not students' self-assessment mark should be included in the summative mark was raised by 7 students (17%), in the second sample discussed in the previous chapter. Five believed self-assessment should be included and two did not. Their responses are presented in Table 7.1.

Respondents' reasons for wanting to exclude self-assessment (n = 2)
Respondent 1: "The major disadvantage for this method is that students assess themselves so you will generally give yourself a good mark."
Respondent 65: "I also realised that I have problems judging my own behaviour – self-assessment is very hard."
Respondents' reasons for wanting to include self-assessment (n = 5)
Respondent 4: "I think there are some students who are over generous in marking their peers, but by including self-assessment they will probably inflate their own total as much as that of their peers."
Respondent 8: "Self assessment is important as individuals may feel that they have put in a lot of effort that other group members may not have been aware of e.g. a lot of research time, therefore self assessment gives the individual a chance to express this."
Respondent 10: ". I feel it is important to be able to include how you believe your own contributions affected the group because only the individual truly knows how much they contributed. This could allow for students to enhance their own grade by rating themselves extremely highly, yet this would be observed by the tutor and extreme ratings could be monitored. Yet individuals who knew that they contributed extra in their own time would be allowed to show this in their own evaluations."
Respondent 12: "It is important in my opinion to assess your own contribution along with the others. Any personal bias will be ironed out by the other members' assessments."
Respondent 16: "Members of a group who are disliked may be falsely assessed without their knowledge and they lose out. That is why it is good to have a self assessment so the tutors can pick up on differences of opinion."

Table 7.1 Students' reasons for wanting to include/exclude self-assessment

Since so few students had commented on self-assessment and the few students who had were divided on the issue, I turned to the literature which included arguments for and against inclusion. Arguments can be divided into two areas: validity and learning benefits for the self-assessor.

Goldfinch (1994), Butcher & Stefani (1995), Healey (1999), Lejk & Wyvill (2001b) and Sharp (2006) are amongst the authors who discuss the validity of including or excluding self-assessment. Goldfinch (1994, p.29) argues for inclusion based on her finding that "over generous students effectively penalised themselves" if self-assessment were not included. (This concern is reflected in the comment of Respondent 4 in Table 7.2.) Goldfinch claims that self-assessment alleviated the problem because overgenerous students would inflate their own score along with everybody else's (Goldfinch 1994, p.30).

Lejk & Wyvill (2001b), on the other hand, recommend excluding self-assessment. They found that high performing students tended to under-rate their own contribution and low-performing students tended to over-rate themselves; thus the inclusion of self-assessment disadvantaged higher performers. Sharp, too, argues that self-assessment should be excluded, citing validity grounds:

> The ratings are intended to measure the strength of each student's contribution. In the case of self-assessments however, personality traits such as self-esteem and reticence may influence the results (Sharp, 2006, p.332).

Butcher & Stefani (1995) and Healey (1999) found that there was no significant difference between self and peer assessed marks.

Boud (1989) and Cowan (1998) stress the importance of self-assessment in terms of the benefits to the self-assessor. Cowan claims that learners who self-assess are more likely to be engaged in deep learning and have a:

> keener appreciation than otherwise of what it is they are trying to do, of how well they are doing it, and thus what they could do next

to improve their performance; they are formatively self-monitoring their process in a directly constructive way to further their learning and development" p.91).

Brown and Knight (1994) claim that unless students have the opportunity to develop self-critical awareness through self-assessment, it is "hard to see how they can benefit from formative assessment. Self-assessment, self-knowledge and formative assessment intertwine" (p.54). This led me to conclude that the act of self-assessing should be included even if the mark were not.

Section 4: Should the allocation of summative marks be confidential?

Although most students in the research reported on in Chapter 6 did not consider whether peer assessment would be anonymous or not, three students in each sample (see pages 95 and 101) did comment specifically on the need for marks to remain confidential in order to avoid conflict and ensure validity. Also, the main reason students rejected option 2 (distribution of a pool of marks) was because the process was *not* confidential.

Although only six students preferred the summative marks allocated to peers for contribution to group work to be confidential and this seemed a sensible course of action, since so few students had offered thoughts on the subject, I consulted the literature to see what previous researchers had written on the subject.

Krause & Popovich (1996. p.142) concur with Johnson (1993, p.1) and Helms and Haynes (1990, p.8) that peer assessment needs to be confidential for maximum effectiveness. This factor was already evident in the responses discussed above. Students who commented on the subject seemed to believe that it is the responsibility of the tutor to protect the privacy and confidentiality of each student's assessment.

Falchikov (1986) and Goldfinch & Raeside (1990) claim that discussion and negotiation of marks amongst students is problematic because students find it difficult to assess each other in such a public way. Topping et al (2000) found that students felt socially uncomfortable when assessing their peers openly. This was reflected in the comments of six respondents reported on pages 95 and 101.

Burnett & Cavaye (1980, p.46) found that, "students feel responsible in making peer assessments but not necessarily comfortable in doing so" and students in Sivan's study (2000, p.196) opted for anonymous assessment forms when the issue of confidentiality was raised. Lejk & Wyvill (2000b) found that secrecy resulted in a greater spread of marks. According to Sharp (2006) this is hardly surprising since students are less likely to want to convey their true opinion of a lazy student in front of them and are more likely to be truthful if their marks remain secret. Sharp (2006, p.331) argues that students' reluctance to be critical in public would result in ineffectiveness in penalising weak contributions and that:

> the real merit of secret ratings is that they are statistically independent – if one student's low opinion of another is not well founded, it will not be reflected in the corresponding ratings of the other students, while if it really is the case that one student had not been pulling his or her weight, the other ratings will reflect this.

For these reasons, I decided that the summative peer assessment process would be confidential but that once the process had been implemented, students would be asked whether or not they thought confidentiality should be maintained in future years.

Section 5: Reflection on the importance of formative assessment

Until this point, discussion has focused on the implementation of a peer assessment process to arrive at a mark for individual contribution to

group work i.e. a summative tool. The aim of this process was to reduce the practice of social loafing, which students identified as a major concern, or to identify social loafers so their lack of contribution could be reflected in their individualised mark for group work. However, given the importance of formative assessment as a learning tool, and students' interest in receiving feedback, it was obvious that the process needed to integrate a formative element.

It is important at this stage to differentiate between the purposes of formative and summative assessment, which are sometimes confused. According to Price et al (2008) and Cooper (2000) summative assessment generates marks for grading purposes and regulates whether students can pass specific boundaries when moving towards accreditation. Formative assessment, on the other hand, is defined by Winne & Butler (1994, p.5740) as:

> information with which a learner can confirm, add to, overwrite, tune or restructure information in memory, whether that information is domain knowledge, meta-cognitive knowledge, beliefs about self and tasks, or cognitive tactics and strategies.

So formative assessment provides feedback that gives students information about how they can progress. To Brown & Glasner (1999) formative assessment primarily involves words and the main purpose is to help students improve, whereas summative assessment is largely numerical and concerned mainly with the end point in terms of evaluative judgements.

Before implementing a formative peer assessment system, I needed to review the literature in order to reforge the links with my espoused theories and to articulate my theories-in-use (Argyris & Schön, 1974) regarding peer assessment. It was important to do this in order to ensure any lines of action were based on a sound rationale rather than faulty intuition.

It is often argued that formative assessment should be an integral part of teaching and learning in HE and that it should be systematically embedded in curriculum practices (see for example Yorke, 2003; Nicol & Macfarlane-Dick, 2007; Price et al 2008). Evidence supports this argument; for example, Butler (1988) found that formative assessment produced more learning than two alternative patterns of assessment: marks only and marks plus feedback.

Hattie & Timperley (2007, p.92) claim that there is considerable evidence to suggest that providing written comments is more effective that providing grades. Black & Wiliam's (1998) substantial review of formative assessment demonstrates its key role in student learning. Their meta-analysis of over 250 studies of formative assessment undertaken since 1988 points to the cognitive advantages and demonstrates that formative feedback results in positive benefits across all content areas, knowledge and skill types and levels of education. Hattie et al (1996) and Price et al (2008) go as far as claiming that of all the interventions aiming to improve student performance, improving aspects of formative feedback has the most impact.

Formative assessment is central to my social constructivist epistemology which sees feedback as a means of bridging what Vygotsky (1978) refers to as the zone of proximal development, the gap between a student's actual and potential level of understanding. Effective feedback provides opportunities to close the gap between current and desired performance (Nicol & Macfarlane-Dick, 2008) and between what is understood and what is aimed to be understood (Hattie & Timperley, 2007).

I believe that peer assessment has other advantages in terms of cognition and metacognition for both the assessor and the recipient of feedback. Skills in norm referencing are developed because assessing students' contribution to the group involves locating one's own performance in relation to the performance of other group members. Topping (1998) points out that providing feedback can lead to

considerable learning not only as a result of having to compare and contrast but also because it involves reviewing, summarising, and communicating.

Cognitive benefits are not the only benefits accrued. Ballantyne et al (2002) outline the pragmatic advantages of involving students in providing feedback to peers. They observe that the unprecedented growth in student numbers has led to the "inability to match resources to the associated marking loads" (p.427). Yorke (2003, p. 283), Ecclestone & Swann (1999), Cowan (2006) and Bloxham & Boyd (2007), amongst others, outline the pressures on higher education that are threatening the use of formative assessment. These can be summarised as:

- Increasing concern with attainment standards, leading to greater emphasis on the (summative) assessment of outcomes.
- Complex tensions between the two roles of tutors as facilitators of learning (formative role) and gatekeepers (summative role).
- Increasing student/staff ratios, and the shrinking unit of resource, leading to a decrease in the attention being given to students as individuals.
- Curricular structure changing in the direction of greater modularisation, resulting in an increased volume of summative assessment and less opportunity for formative feedback.
- The demands placed on academic staff, which include the need to be research active and generate funding, in addition to teaching.

One of the key issues resulting from the pressures facing tutors is how to ensure that students receive *enough* useful feedback. Peer assessment makes pragmatic sense because it increases the amount of feedback students receive. In the case of the present research, students are the only people in a position to provide each other with feedback on contribution to group work in terms of their weaknesses and strengths in the skills involved.

Since the tutor is not in a position to provide feedback on group working skills, because meetings take place outside taught sessions, it is important that students take on this responsibility. Not only are they better placed to do so, but the language they use is sometimes more easily understood by recipients. This is because academics often use a language of assessment that students do not understand (Millar, 2008; Smart & Dixon, 2002; Bryan, 2004).

Ballantyne et al (2002, p.427) claim that peer assessment benefits both assessee and assessor; the former receives more feedback than would be possible if one tutor were the sole source of feedback and the latter has the cognitively challenging task of engaging with new knowledge and developing skills in assessment.

Peer assessment, whether formative or summative can also have an inclusive function: it can help socialise students into a community of assessment practice. It can afford greater insight into assessment by demystifying the tutor's tacit knowledge for the students (Klenowski, 1995; O'Donovan et al, 2008), thereby enabling students to appreciate why and how marks are awarded (Brindley and Scoffield, 1998). It can also provide students with a better understanding of what is required to achieve a particular standard and what academic staff are looking for when conducting assessment (Falchikov, 1995; Hanrahan & Isaacs, 2001; Race, 1998; Ballantyne et al 2002, Bloxham & Boyd, 2007). If students are to acquire the tacit knowledge acquired by tutors, they need to engage with the assessment process from the start of their academic careers.

Providing feedback can also help students acquire skills relevant to the work place. Brown & Knight (1994, p.61), for example, claim that it is important that students are:

> exposed to situations which require them to respond sensitively and perceptively to peers' work. As in employment, those whom

> we judge now are ones with whom we shall be working again. Much can be learned about leading teams and working within a team through the experience of peer assessment.

Giving peer feedback involves learning how to give constructive criticism diplomatically, a skill of key relevance in adult life.

Feedback about group working skills is vital if students are to improve in this area. Although using peer assessment to provide a summative mark may go some way to alleviating the problem of social loafing, the aim of formative assessment is to improve learning while it is happening in order to maximise success rather than merely determining success or failure only after the event (Topping, 1998). It has a central role to play in ensuring an effective learning experience.

Section 6: How could a formative element be incorporated so that students receive the benefit of feedback to facilitate improvement of their group working skills?

Having articulated my reasons for wanting to include formative feedback, the next question to address was how students could provide this feedback to others in their group. Although Strachan & Wilcox (1996) acknowledge the need to develop some way to pass on the evaluative comments of peers, they do not suggest how this might be implemented. Krause and Popovich (1996) suggest that students write additional comments on the feedback sheets on each person and this should then be typed up by the tutor so that the handwriting would not reveal the student's identity. They advise: "To encourage the student evaluators to be objective in their assessments and to help students value the assessments of their peers, the instructor must ensure that the feedback to the student is presented in an anonymous, constructive manner" (1996, p.142). Strachan & Wilcox (1996) also suggest that formative comments remain anonymous.

However, I felt uneasy about keeping the formative assessment element anonymous for several reasons. Firstly, if students are to engage in the assessment process and be integrated into an assessment community, they need to be encouraged to take responsibility for their actions, and that should include the feedback they provide to peers. Tutors' feedback is not anonymous. Secondly, since identifiability is a key factor in ensuring students take responsibility for their actions, I felt that those providing feedback should not be encouraged to hide behind a cloak of anonymity. Thirdly, if students know their feedback can be traced back to them, they might be more likely to take more effort when formulating it. Fourthly, if students know who has written the feedback, they can seek clarification should they need to do so. Fifthly, summative assessment involves deciding on a mark in terms of a number of agreed criteria but does not involve students in having to think about how to *articulate* and *justify* their praise and criticism in a sensitive way.

Due to these reasons I decided instead that students would type up their feedback to each other member of their group in advance and bring it to the final session. An example of the form used can be found at Appendix 3. I decided to evaluate students' responses to the process, so that the decision could be reviewed in the light of student feedback.

Conclusion

This chapter has discussed a number of questions. The first question was: should peer assessment of contribution to group effort be based on one holistic figure or on a number of pre-specified criteria? The literature suggested that criterion-referencing was the better option for several reasons which were discussed. However, I wanted my students to be given the opportunity to make their own choice. The findings concerning students' choice between criterion-referenced and holistic peer assessment are reported in the next chapter.

The second question was: If students chose criterion referencing, who should decide on the criteria to be used? The literature was clear that if students were to be given the opportunity to be integrated into an assessment community, take more responsibility for their learning and acquire tacit understanding of assessment, then they should take on this role. I decided to evaluate to what extent students agreed with this if they chose to implement criterion-referencing assessment.

The third question was: should self-assessment be included or excluded? I decided that students should assess their own contribution due to the advantages accorded in terms of learning but that the mark would not be counted when calculating their overall mark in case high performing students under-rated their own contribution and low-performing students over-rated themselves (a tendency reported in the literature).

The fourth question involved whether the allocation of summative marks should be confidential. I decided to follow the guidance in the literature and ensure that summative marks remain confidential but to evaluate students' opinions of the process.

The fifth question looked at how a formative element could be incorporated so that students received the benefit of peer feedback to facilitate improvement of their group working skills. I decided that students' feedback should be given face-to-face in the final session and that an evaluation of students' responses to the process would inform my future practice.

The next chapter reports the findings concerning students' choices between criterion-referenced and holistic peer assessment.

Chapter 8: Assessment decisions and implementation

Chapter 7 reported on a literature review carried out in order to ascertain the advantages and disadvantages of holistic and criterion-referenced peer assessment. Although an analysis of the arguments pointed towards the latter being a more effective system for a number of reasons, which were discussed, I wanted my students to be given the opportunity to make their own choice.

In Section 1 of this chapter students' responses to the following question are reported:

- Would criterion-referencing or holistic peer assessment be a more valid method to use when evaluating the contribution group members made to a group project?

Section 2 describes the evolving process for deciding on the criteria for assessment and explains how assessment was implemented.

Section 1: Criterion-referenced or holistic assessment?

This section briefly describes the method used before reporting and discussing the findings of research relating to the above research question.

Method

Three groups of second and third year education students (n = 99) were asked to consider whether criterion-referencing or holistic peer assessment would be more valid when evaluating the contribution group members made to a group project (Questionnaire 4). (The notion of validity was discussed in some depth with each group beforehand.) It was

made clear to students that the choice they made would be implemented in the module. The differences between the two forms of assessment were explained on a handout (summarised in Table 8.1) and discussed in class.

Criterion-referenced (CR) peer assessment involves summative assessment against agreed criteria. So, for example, you would assess each group member against a number of agreed criteria (such as contributions to discussions, attendance at meetings, etc.) negotiated at the beginning of the module.

Holistic (H) peer assessment would involve giving each group member a grade/ mark based on an intuitive, general 'gut' feeling about what each member deserves. It would involve one mark only and there would be no criteria on which to base your judgement.

Table 8.1 Excerpt from handout of differences between criterion-referenced and holistic assessment discussed with three groups of students.

Space was provided for students to give reasons for their choice. Students' responses were transcribed and analysed using the method described in Chapter 3.

Discussion of findings

Ninety-two students completed Questionnaire 4 in which they were asked to decide which of the two options they considered more valid and why (response rate 93%). The results are shown in Figure 8.1. Eighty-six out of the 92 respondents (93%) reported that CR peer assessment would be more valid. Five students felt that the holistic method would be more valid (5%). One student felt that peer assessment should incorporate both elements.

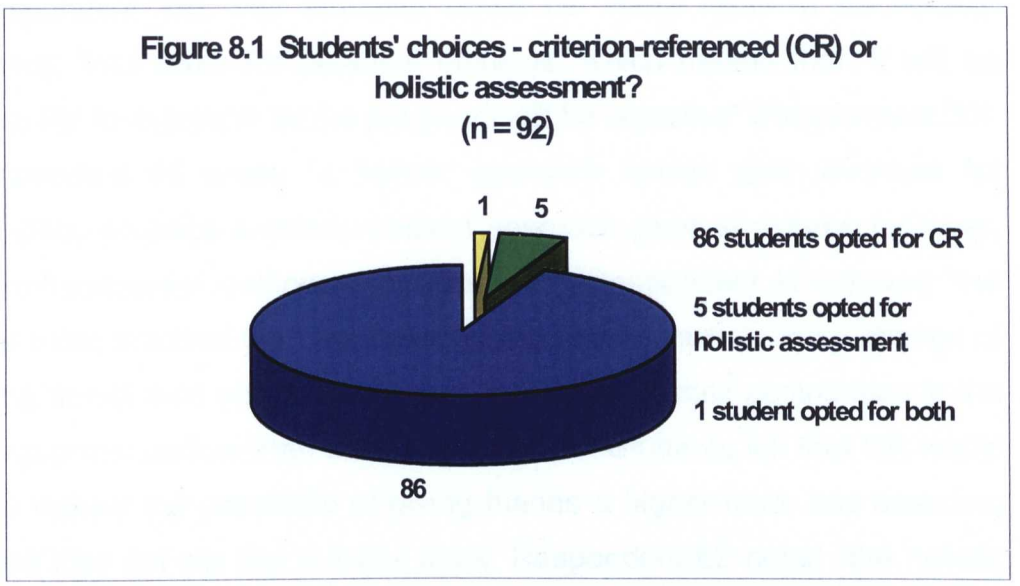


Figure 8.1 Students' choices of criterion-referenced versus holistic assessment.

Five themes emerged from students' responses:

- Holistic assessment is open to 'bias'/'prejudice' while CR assessment is fairer because it is less 'personal', more 'objective'.
- CR assessment facilitates differentiation of skills and contributions. Weaknesses and strengths are acknowledged and input is more likely to be valued.
- CR assessment results in standardisation; it allows for more 'structure', is more 'accurate', 'specific', 'detailed', 'clear', 'quantifiable' and 'measurable' which are seen as positive attributes by students. This makes assessing easier.
- Because CR is specific, students are clearer as to what is expected of them in terms of behaviour and skills when working as part of a group.
- CR assessment opens up the opportunity to receive feedback on each of the criteria included.

Forty-four students (48%) felt that holistic assessment would be open to 'bias' or 'prejudice' and that CR assessment was fairer because it was less personal and more objective. For example, students felt CR would reduce the impact of the assessor's "opinions towards the assessed"

(Respondent 13), that students would be "more likely to be honest" leaving "less room for personal opinions" which means that "it will be more fair to everyone as the judgment will be objective" (Respondent 39). Respondent 48 wrote: "a holistic approach leaves open avenues for prejudice whereas a criterion-based approach gives structures for fairer, more transparent methods of assessment." Respondent 47 believed "it is also more practical than the holistic [method] as there is more chance of being non-biased and assessing the individual on their contribution to the group project rather than them as a person." Students felt that CR would help reduce the possibility of giving friends a higher mark and awarding those they did not like a lower mark. Respondent 62 noted "the holistic way is problematic if you have to assess a friend of yours or a person you generally dislike – you won't assess objectively" or as Respondent 92 points out, "The holistic method could be a very unfair method for assessing group members. Conflict amongst members could result in biases and prejudices whereby individuals are not fully recognised for their strengths."

Thirty-two students (35%) felt that CR assessment facilitates the differentiation of skills and contributions group members make and that these can then be acknowledged. Respondent 11 stated: "it's a fairer system. The student might have excelled on one area but not have been so effective in another area" while Respondent 12 noted "it will show strengths and weaknesses in different areas so they can receive a more fair mark" and Respondent 22 commented, "[it is] a fairer marking system as it identifies people's strengths and weaknesses and where exactly their contributions were." Respondent 52 suggested that "assessment should be based on a number of different criteria which could be task and relationship orientated. This would give peers a chance to comment on a broad range of criteria." Students seemed to like the fact that different sorts of skills and abilities could be recognised, for example the person who "may be quiet at meetings and not contribute but does a lot of research" (Respondent 72) so "the quiet and diligent people [....] may not be rewarded" (Respondent 74) if a holistic method of assessment were to

be adopted. This in turn means that people are more likely to feel valued: "contributions are valued more" (Respondent 32) and "your contributions are more likely to be valued" (Respondent 37).

Forty-three students (47%) felt that the standardisation facilitated by CR made it more valid, more 'structured', more 'accurate', 'specific', 'detailed', 'clear', 'comprehensive', 'quantifiable' or 'measurable.' Students thought that it was more valid "because everyone would be judged on the same guidelines" (Respondent 4); "it's actually based on something whereas holistic assessment is just a gut feeling not based on anything concrete" (Respondent 8). According to Respondent 21 CR is more valid because "everyone is judged from the same point of reference. If there were no indications then scores would be all over the place. Guidelines for how to assess other people need to be fair". For Respondent 53 "holistic is too general and not specific enough whereas CR is more structured and thorough and is clearer for the student to understand". Students valued clarity, for example Respondent 61 said that CR was more valid "because using criteria will enable the assessment process to be explicit. Discussion can take place among group members to decide exactly what the meaning of the criteria are and what is being assessed". Seven students commented on the fact that they were more likely to know what was expected of them with a CR method of peer assessment, as Respondents 82, 92, 9 and 46 noted respectively: "As a group you must agree on the criteria so therefore know how you should perform" and "because it allows the whole group to initially agree on the criteria, in so doing the criteria then act as the group guidelines or standards"; this means "they know from the outset what the expectation is" and "can work towards achieving the criteria".

Thirteen students (15%) felt that CR assessment opened up the opportunity to receive constructive feedback. For example, Respondent 25 wrote " feedback is useful; knowing where you went wrong, etc. Group evaluation of your role is a positive idea. Good to know where you stand". Respondent 47 felt that as CR is more structured it leads to "better

feedback" and "would enable someone to see in what area their group skills are lacking" (Respondent 52) which would be useful for " future work" (Respondents 65 and 68) and give the recipient the opportunity to "improve" and "make changes" (Respondent 91). Respondent 55 noted "that peer can go away recognising what they're good at and what they need to improve" whereas Respondent 62 asserted, "the holistic way doesn't provide feedback which helps you to work on certain skills."

Respondent 19 was one of five students who felt that an holistic method was more valid because:

> there might be things that happen that affect the group that are not listed in any of the theories. Also just because someone hasn't pulled their weight on all/most of the criteria it doesn't mean they weren't valued.

Respondents 64, 73 and 15, although concluding that CR was more valid, agreed with the first part of respondent 19's criticism. Respondent 64 commented that one disadvantage of CR "is that the pre-set criteria may have nothing to do with the real problems that come up so the criteria may be irrelevant" while Respondent 73 stated "sometimes set criteria may turn out to be irrelevant" and Respondent 15 wrote "something could be missing off the criteria."

This criticism is echoed by two more of the five students whose preference lay with holistic assessment:

> [it] looks at the person as a whole rather than judging them by criteria which may not apply to them. When in a group you experience them as a whole not as a set of criteria. (Respondent 70)

> In the criteria method, some things that one of the group members has done may not fall into any of the criteria so it will go unrecognised. CR is too structured and does not cover everything. (Respondent 80)

A solution is proposed by Respondent 78 who suggests that room should be made "for a couple of lines for members to add additional comments as the criteria may not cover all aspects considered important." However, the tutor's role as an "expert guide" (McDowell & Harman, 2008, p.4) can help ensure that when students are negotiating criteria at the beginning of the module, they have considered all the relevant skills.

Gut feelings are not rejected as lacking objectivity by three students who nevertheless opted for CR. "It is possible to just have a gut feeling that a member has not earned their grade (Respondent 16); "sometimes a first impression can be correct" (Respondent 73) and "the first impression could be right and the disadvantage of having too many criteria is that you can't see the wood for the trees" (Respondent 64). Respondent 86 claimed:

> I am an objective person – in other words it is much easier for me to evaluate general contribution rather than each aspect of it. (Respondent 86)

However, none of these students considered the difficulty of articulating gut feeling.

To summarise, overall CR was seen as 'fairer' (the word fair was mentioned 29 times), less open to bias, more discriminating, allowing for more standardisation, resulting in clearer expectations and opened up the opportunity for constructive feedback based on the criteria involved.

These findings contradict the findings of Lejk and Wyvill (2001a) who report on the results of their research comparing a holistic and category-based approach to peer assessment of contributions to a group project. Their research involved 155 Year 2 students in groups of three to six over a four week assessment period at the end of a twelve week semester. The category-based approach was based on the work of Falchikov. Their 2002 paper concludes that student attitudes were more supportive of the holistic rather than the category-based approach.

However:

- Their assertion that a holistic approach is better than a category-based approach is based on a comparison of two groups of students (according to this paper). However, the semester 2 students "appeared to be a more committed group of students" (their attendance was much better) (p.576) who may have been equally supportive of the category-based method had they had the opportunity to experience it.
- The assessment criteria were not developed by the students themselves so there is no evidence that they understood them.
- Their 2002 paper contradicts information reported in their 2001a paper which refers to 155 students and states :"In Week 11 the students submitted two peer assessment sheets, one being a category-based sheet and the other being a holistic assessment." In their 2002 paper it says that, in fact, these were two different groups of students and that each group of students undertook either a category-based or a holistic method.) Also their 2002b paper reports on 172 semester 1 students who undertook two sorts of category-based assessment: in week 11 they undertook "secret" category based assessment (i.e. they did not tell their peers what marks they had given them) and in week 12 they completed an "agreed" category-based assessment sheet (i.e. the marks were discussed among the group and agreed on). So maybe this group were less supportive of the category-based method because a) they saw the difference between the marks awarded depending on whether they were done secretly or after discussion, or b) the process of discussing individual marks could have affected them negatively. (Falchikov, for example, reports that students admit to not liking awarding marks to friends, so it is easy to imagine that they find the whole process of discussion distasteful.)

Since 93% of the students who completed questionnaires in my study felt that CR was a more valid assessment method, I decided to adopt this approach.

Section 2: A description of the process for deciding on the criteria for summative peer assessment of contribution to the group during the module.

It was established in Chapter 7 that students' ownership of the criteria used for assessment was of fundamental importance (Claxton, 1995; Hall, 1995; Klenowski, 1995; Dochy & McDowell, 1997; Pond and ul-Haq, 1997; Falchikov, 1998; Stefani, 1998; Ross and Rolheiser, 2003). In addition I felt that this feeling of ownership would improve group cohesiveness which was identified in Chapter 5 as an important factor in discouraging social loafing by increasing the value of the group to individual group members.

A session was held at the beginning of each module (once assessment tasks had been explained and discussed) to establish the criteria for peer assessment of contribution to group effort. Race's (1998) process for agreeing criteria for a presentation was adapted and the procedure in Table 8.2 below introduced.

1. After the group assessment tasks were explained, students were asked to write down some key words about "what makes a really good group member".
2. Students were then divided into groups of four to discuss their ideas.
3. The class then fed back their ideas so that I could ensure they were describing assessable skills or behaviours rather than personality traits.
4. Each group member then made a short list of "the five most important behaviours/skills of a good group member" generated in their group.
5. New groups were then formulated a number of times and the process repeated with individuals adding to their list as necessary.
6. After 2 – 4 changes of groups (depending on time available) each group then appointed a scribe and came to an agreement on a short list of "the five most important behaviours/skills of a good group member" (Questionnaire 6).
7. After a class discussion and a subsequent opportunity for groups to revise their short-lists, the lists were collected and the information

analysed to determine the main skills/behaviours identified by students.

8. These findings were presented to students the following week. Only skills/behaviours included by more than one group were included.
9. Grade descriptors were then produced following discussion with students. These were formulated as a grid and given to students in a pilot study but were replaced by a Likert scale after feedback from students.
10. At the end of the module students used a form to complete summative self and peer assessment in class. (See Appendix 2 for an example). Students sat away from all other members of their group so they could concentrate on writing without worrying about peers trying to read their marks.

Table 8.2 Procedure for students to decide on criteria on which summative marks are to be based

Table 8.3 shows the skills/behaviours identified by one cohort of 30 groups of three – five students (n = 124).

Skills/behaviours identified	Number of groups choosing this skill/behaviour
Quality of contribution (ideas and written work)	28
Respect for others (listening to and respecting others' opinions and ideas)	27
Time management (meeting deadlines; being punctual)	26
Cooperation (supporting others; willingness to compromise; flexibility)	29
Commitment (attendance; preparation; effort; enthusiasm)	28
Reliability (doing what they say they will do)	30
Organisation skills (organising others, meetings, content, resources)	28

Table 8.3 Skills/behaviours chosen by thirty groups of students as criteria for peer assessment on one module

A form was then compiled which included the skills/behaviours identified (see Appendix 2 for an example).

Conclusion

Students in research carried out in the first cycle of this study (reported in Chapter 4) identified the practice of awarding group members the same mark irrespective of individual contribution as a major problem in assessed group work. Students expressed the desire to identify individual contributions to assessed group work in order to deter students from social loafing and identify students who did not contribute.

In the second cycle of research, students on a number of modules were actively encouraged to make decisions about how to alleviate the problem of social loafing. Seventy-seven out of 97 second and third year students (79%) chose to address social loafing by using a peer assessment process to identify the contribution of each group member and so encourage peers to work harder while making it possible to identify those who did not pull their weight.

These students elected to base assessment on a list of behaviours which they devised collectively in consultation with me. The resultant summative mark remained confidential, but a formative element was included in the process which involved students discussing written feedback each gave and received from other members of the group in the final class.

The next chapter evaluates the summative and formative process from the students' point of view.

Chapter 9: Students' perceptions of the summative and formative peer assessment process

This chapter discusses students' perceptions of the summative and formative peer assessment process described in the previous chapter. Section 1 explores students' feelings about being involved in the process of setting criteria. Section 2 reports students' responses to both formative and summative peer assessment and explores whether they would prefer the process to be anonymous or not. Section 3 discusses to what extent peer assessment of contribution alleviates the problem of social loafing. Section 4 contains a summary and discussion.

Section 1: How did students feel about being involved in setting criteria?

Method and discussion of findings

Students on two different first year modules were asked what they thought about being involved in the process of deciding criteria for assessing each other's contribution to group work. The students completed a short questionnaire (Questionnaire 7), which was anonymous. Seventy-five out of 91 students in the first group and 104 students out of 129 students in the second group completed the questionnaire (response rate 82% and 81% respectively). Ninety-three per cent (i.e. 70 out of 75) in the first group and 82% (i.e. 85 out of 104) reported that they thought it was a good idea to be involved in setting criteria for assessment.

In Chapter 7 various reasons why students should be involved in setting assessment criteria were discussed. The arguments included giving students more ownership of assessment to encourage more responsibility for their own learning and to socialise them into an assessment community. Analysis of students' responses supported these arguments.

Students appreciated the inclusive nature of being involved in setting criteria, for example Respondent 3 commented: "It enables you to discuss what you believe to be important" as opposed to "a lecturer's point of view," whilst Respondents 68, 10 and 48 wrote respectively: "We get to be involved"; "it's nice to be involved"; "you feel more involved."

Appreciation at having their voices heard was a theme which emerged from comments made by other students such as Respondent 25: "This way what the students think is important is taken into consideration", Respondent 29: "Allowing people to come up with their own criteria creates a more inclusive and productive atmosphere", Respondent 48: "It's our hard work so it's nice to be involved in setting criteria." Respondent 73's comment manages to communicate incredulity when he/she writes: "[the criteria] were actually taken into account and used."

The opportunity to be involved in setting criteria seemed to increase students' ownership of the process to the extent that Respondent 75 suggested: "Perhaps each group should decide instead of the whole class?" This confirms research undertaken by Searby & Ewers (1997), Cowan (1998), Pond and ul-Haq (1997), Falchikov (1998) and Sivan (2000) who found that students' involvement in setting the criteria was an essential strategy to maximise the potential of peer assessment for developing students' sense of ownership and control and for allowing them to exercise responsibility for their learning.

Claxton (1995), Peters (1996) and Pain et al (1996) suggest that negotiating assessment criteria increases students' acceptance of them and promotes student reflection. It also takes some of the mystique out of the marking process (see also Mowl & Pain, 1995: Ballantyne et al, 2002; Brindley & Scoffield, 1998). This certainly seemed to be the case with the group of students involved in this research. Respondent 23 wrote, for example: "It works so well because it gets students thinking about how they are being assessed," while Respondent 74 claimed: "As I have contributed to this [i.e. designing the criteria] it makes me think more

deeply about how I can achieve a good mark." Many students expressed the belief that being involved in creating the criteria made the whole process more transparent and increased their achievement. For example, Respondent 35 stated: "[I had a] clear understanding of what we had to do to score good marks", Respondent 62 wrote, " it clarifies the whole process", Respondent 30 reported, "It gives you more understanding and makes the whole idea of assessment less daunting" and Respondent 55 concluded, "you can work towards something and know what is expected of you. You can appreciate your grade if you know what you are being marked against." These comments highlight the worrying possibility that students do not normally understand assessment criteria (confirmed, for example by Millar, 2008). If this is indeed the case, then Ross and Rolheiser (2003, p.130), Searby & Ewers (1997), Dochy & McDowell (1997), Hall (1995), Boud (1989) and Bryan (2004) are justified in stressing the importance of involving students in the setting of criteria in order to help them develop a shared language of assessment. This in turn increases the likelihood of students acquiring the 'tacit' understanding referred to by Rust et al (2003) and being socialised into the assessment community proposed by Ecclestone and Swann (1999) and Price et al (2008).

Students on these two modules together with the second sample of students discussed in Chapter 6 were not asked to evaluate the actual summative and formative peer assessment process. The reason for this was that I felt that I had already requested enough of their time in asking them to complete questionnaires. For this reason, students on another module in which the same process was introduced were asked to evaluate the process of summative and formative peer assessment. This is reported in the next section.

Section 2: Students' evaluation of the summative and formative peer assessment process

Method

One hundred and thirty-eight undergraduates on a first year module on Human Communication working in groups of three to five were asked to agree on assessment criteria at the beginning of the module. At the end of the module students assessed each other and themselves on their contributions to group assignments and completed two forms. One form (Appendix 2), which involved allocating a summative mark based on the agreed criteria was completed during the final session and was confidential. The second form (see Appendix 3) was completed prior to the final session and involved each group member compiling formative feedback for each other member of the group, based on the criteria agreed at the beginning of the module. Students gave each other these feedback sheets in class, discussed the content and had the opportunity to ask their peers for clarification about what they had written. At the end of the session, students were asked to complete Questionnaire 8 (see Appendix 4) which contained a number of questions about their experience of the summative peer assessment process together with giving and receiving feedback (formative assessment). The questionnaire was completed in class after students had undertaken the feedback exercise and was voluntary. It was attached to the back of a module evaluation form (which is anonymous) so students did not feel embarrassed if they did not want to complete it.

The data were analysed using the method described in Chapter 3. Two samples of text were assessed to establish inter-rater reliability with two different colleagues and a group of twenty-five second year students on a Research Methods module.

One hundred and eight of the 138 students completed the questionnaire (response rate 78%). Not all students gave answers to all questions. Students were asked whether or not they thought it was a good idea to receive feedback from other group members on the contribution they had made and their group working skills. Ninety-seven students (90%) replied that it was, 8 (7%) felt that it was not and 3 (3%) were unsure. Figure 9.1 illustrates the findings.

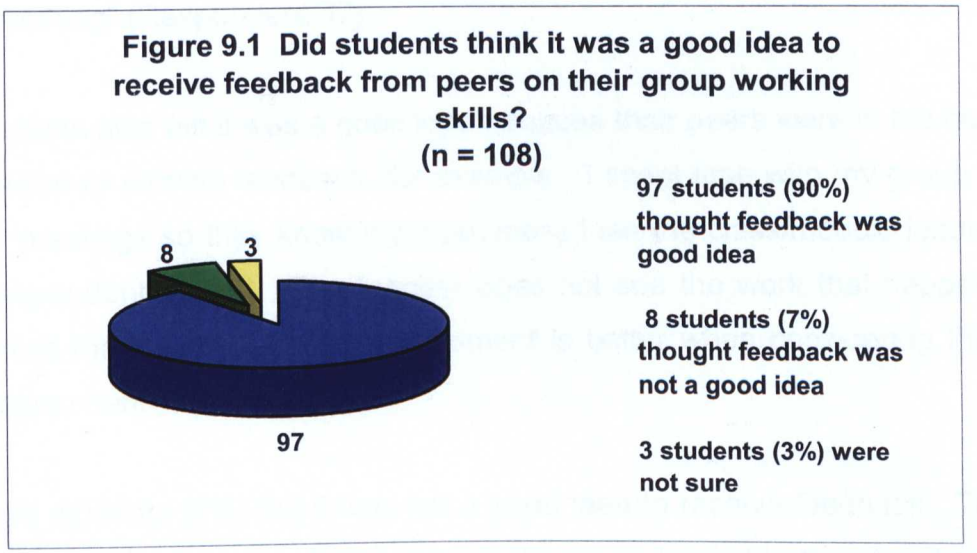


Figure 9.1 Did students think it was a good idea to receive feedback from peers on their group working skills?

Students gave two main reasons for stating that formative peer assessment was a good idea. Firstly, they saw it as an effective learning tool since it enabled them to identify in what way they needed to improve their group working skills. Eighty-two per cent of students (n = 80) welcomed the idea of *critical* feedback. Respondent 10 explained: "If you know what you've got to improve, you will succeed better when working in a group again." As Respondent 80 pointed out: "Otherwise you don't know where you're going wrong." Respondent 35 stated: "It is especially good for me as group leader to see what I can do better in the future." Students were also keen to learn what others thought of them, for example, Respondent 29 wrote: "I am particularly glad we did the

feedback task as I think it is useful to me to see what others thought of the way I work." Respondent 34 stated: "It's interesting to learn what people think of you" and Respondent 103 claimed: "I like to know and am interested in how other people interpret me."

Twenty per cent of students (n = 19) reported that they thought it was a good idea to receive information about their group work strengths i.e. positive aspects of their group work skills "because it helps to know what you are good at" (Respondent 42) and "because you can see what you've done well over the module [which] makes you feel you've achieved something" (Respondent 97).

Students also felt it was a good idea because their peers were in the best position to provide feedback, for example: "I spent time with my group in our meetings so they know my input more than the class/module leader" (Respondent 5) and "The teacher does not see the work that happens behind the scenes so peer assessment is better when considering this" (Respondent 12).

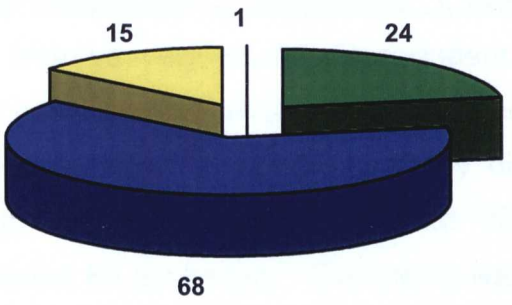
Eight students (7%) felt it was not a good idea to receive feedback. The main reason given was that peers would not be honest in their feedback, for example Respondent 74 stated: "It is good to see what people think but most of them may be more nice than they should".

Three students were not sure whether peer feedback was a good idea or not. Respondent 81 gave no reason for his/her answer. Respondent 53 felt that "although it helps to know where you need to improve, it also means you are on edge about group assessment." Respondent 48 said "honestly, it is my first time assessing other people so I'm not sure if it is good or not."

Students were then asked to what extent they found the feedback from their peers beneficial. Eighty-five per cent of respondents reported that it

was "very" or "quite" beneficial," (n = 92), 15 (14%) felt it was "not very beneficial" and one student found it "not at all beneficial" (see Figure 9.2).

Figure 9.2 To what extent did students find feedback from peers beneficial? (n = 108)



24 students (22%) found feedback very beneficial

68 students (63%) found feedback quite beneficial

15 students (14%) found feedback not very beneficial

1 student found feedback not at all beneficial

Figure 9.2 The extent to which students found feedback from peers beneficial

Of the 92 students (85%) who reported that they found feedback from peers "very" or "quite" beneficial, the main reason given was that the feedback enabled them to identify their strengths and/or weaknesses, in other words it fulfilled the function of providing information to help them learn about themselves. Sixty-seven students (62%) commented that they had learned more about themselves, for example Respondent 18 stated: "They brought things to my attention I hadn't thought about"; Respondent 23 reported: "It showed where my strengths lie but also pointed out weaknesses that went over my head that were never addressed during a meeting"; and Respondent 102 wrote: "It helped point out what I need to improve on but also what I bring to the group which I was not necessarily aware of before."

Fifteen students (14%) reported that they found the feedback "not very beneficial" and one student found the feedback "not at all beneficial" (but did not give a reason why). Eleven of the fifteen students in the former category gave lack of validity as the reason they did not find it very beneficial, i.e. the information provided was too vague, as did 16 of the 68 students who found the feedback "quite beneficial." For example the feedback Respondent 16 received "lacked examples"; Respondent 101,

who received mostly positive feedback questioned: "is this total truth?" and Respondent 74 admitted: "I thought they were too nice to me. I didn't think that I contributed enough on the poster." Although students welcomed positive feedback, they also felt that the balance between positive and negative comments was skewed in favour of the former. For example Respondent 3 complained: "It was largely positive so didn't help me to improve very much", Respondent 8 reported he/she received: "mostly positives and I'm sure there were some things I could improve on" while Respondent 72 lamented: "They didn't give me any criticism or ideas to improve myself. I wanted them to be more critical" and Respondent 83 confirmed: "Everything was fairly positive. You don't gain anything from that."

Twelve students (11%) felt that they did not learn anything new from the feedback, for example Respondent 35 claimed: "I knew intrinsically what I needed to do anyway," while Respondent 51 stated: "It was quite similar to what I predicted" and Respondent 81 asserted: "I know what mistakes I make." Other students found this confirmation reassuring, for example Respondent 68 who commented: "They picked out my weaknesses which I was already aware of. I suppose it was reassuring."

I was particularly interested in students' perceptions of the experience of providing feedback since there is evidence in the literature (for example McWhaw et al, 2003) that students are uncomfortable with both formative and summative peer assessment. Students were asked how they felt about the feedback process.

Of the 108 students who completed the questionnaire, five students did not answer this question. Sixty-six of the 103 students (67%) expressed positive feelings about providing feedback. Words used to describe the experience included "good", "brilliant" or "excellent" (22 students) for example: "This was a brilliant opportunity to see how well you work with a team" (Respondent 104); "I think it is an excellent idea. I understand its place and it will be invaluable for future reference on what I need to

improve on" (Respondent 108); "It was good as it helped me to think back and look at where I personally went wrong" (Respondent 4); "I thought it was a good idea as it meant that we could see from the other group members' perspectives what our performance was like" (Respondent 12).

For 21 students the process of writing feedback evoked other positive feelings, for example, confidence, enjoyment, and interest for example, Respondents 24 and 39: "felt (very) confident ...." "Useful", "invaluable", "beneficial", "helpful" were also used, for example Respondent 75 found the process "useful for them and to me because I could think about their actual participation" and Respondent 97 stated, "It was useful as it did make me think about what skills you need to work in a group.....good to learn to write it constructively." What is noticeable from the comments of the students above is that the process of having to articulate feedback resulted in meaning-making. Students seemed to engage in reflection-on-action (Schön, 1991; Cowan 1998).

However, despite the fact that the majority of students reported feeling positive about the experience, 33 (32%) students used the words 'hard', 'difficult' or 'not easy' on a total of 42 occasions to describe their feelings about the task of providing written feedback to their peers. This accords with research carried out by Falchikov (1995), McDowell (1995) and Mowl and Pain (1995) who report that the majority of their students found peer assessment difficult. Six students also reported feeling "uncomfortable" or "awkward" and five students wrote that they felt "anxious", "nervous", "under pressure", "emotional" or "scared."

Students used the above words to describe the act of providing criticism. There were four main reasons students associated providing criticism with negative feelings. Some students (n=14) found it difficult to find anything negative to say. Although it is possible that these students found themselves in groups in which everything ran smoothly, Tuckman (1965) argues that disagreement or "storming" is a necessary stage in group

development, so perhaps students in these groups actively avoided conflict or did not get to know each other well. Respondent 22 wrote: " I found it hard as we all worked so well together and got on well. I tried to find stuff they could improve on but it was too hard" and Respondent 52 added: "[It] was difficult as I feel the entire group worked really well together and to be honest I found it very hard to think of negative points."

Other students (n =11) found it difficult to *articulate* criticism. Respondent 4, for example wrote: "I found it difficult trying to word any criticisms I was making." Respondent 63 reported:

> In reality it's not easy to do: honestly I don't know how to assess my group mates because sometimes you just feel like that but when you are asked to give reasons or examples, you just can't do it.

Eight students seemed reluctant to hurt others' feelings by criticising them. Respondent 10, for example, found it difficult "to find negative stuff to say as I didn't want to hurt anyone's feelings" while Respondent 34 "found this hard and quite emotional because I didn't want to criticise my group as we all worked hard."

Seven students felt that it was difficult to provide negative feedback *face-to-face*. Respondent 38, for example, found it:

> difficult, as I knew they would be reading it in my presence and they'd know I had written it. I found it hard to be critical under these circumstances.

Respondent 83 concurred:

> It was hard as it wasn't anonymous so you feel bad if you give constructive criticism to them as you can see instant reaction after they read the comments.

Two students admitted to getting round the awkwardness of providing negative feedback by avoiding giving it in the first place, for example

Respondent 14 conceded: "I didn't feel like I could be 100% honest and therefore left out some information to spare the person's feelings."

Five students complained about the time and effort involved but three of these admitted the benefits. Respondent 3, for example, thought that "though time-consuming, it was a useful activity and I appreciated receiving my feedback" while Respondent 45 admitted "It felt like a bit of an annoying task when we have lots of other essays due in, however, writing it showed it was beneficial to me in understanding our group's dynamics and everyone's strengths and weaknesses."

Students were asked whether or not they thought the process of giving feedback should become anonymous. Responses are shown in Figure 9.3. Fifty-five students (51%) reported that they would have preferred to give their feedback anonymously and 53 (49%) said they preferred the current system of providing written feedback face-to-face.

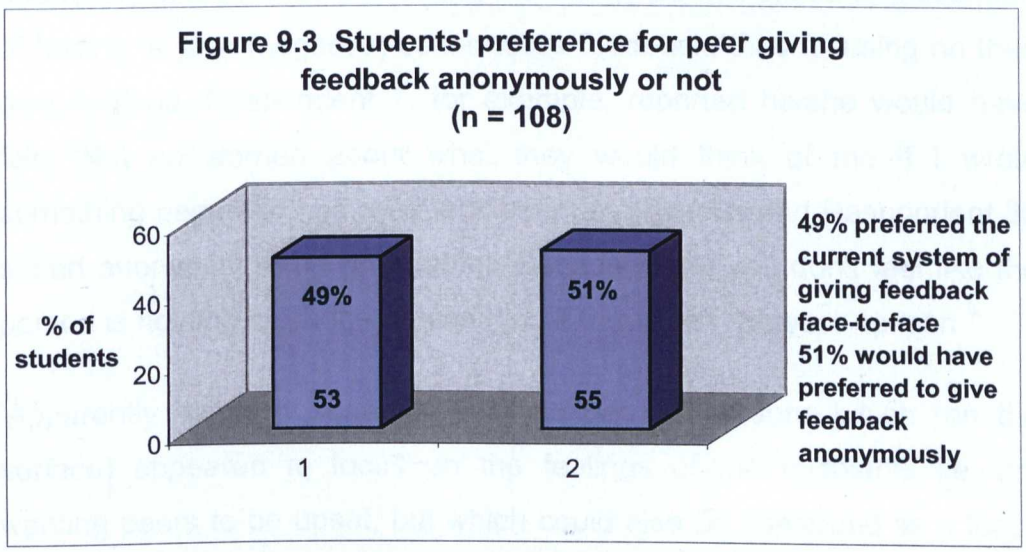


Figure 9.3 Students' preferences for giving feedback anonymously or not

The reasons given by the 55 students (51%) who would have preferred to have given anonymous feedback fell into three categories: egocentric reasons, apparently altruistic reasons and validity reasons, illustrated in Figure 9.4.

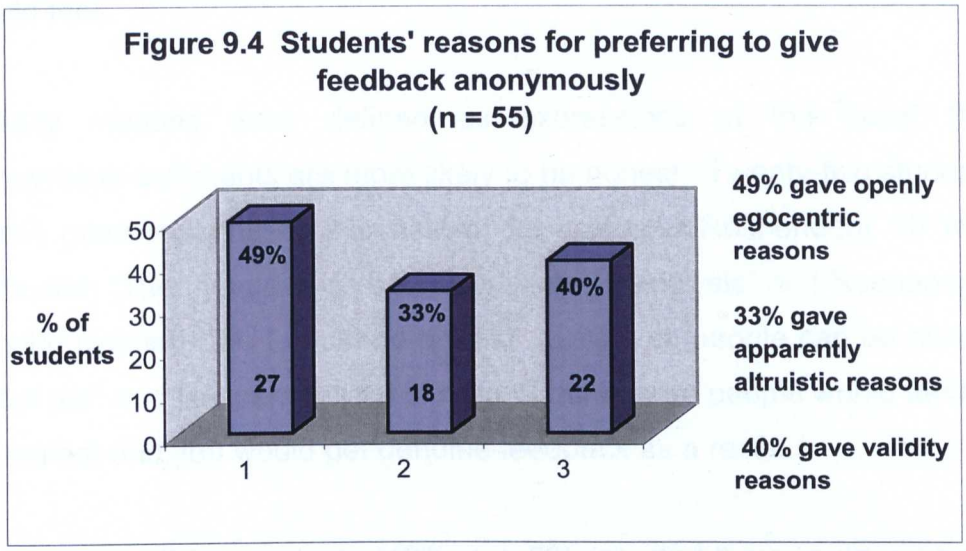


Figure 9.4 Students' reasons for preferring to give feedback anonymously. (N.B. some students gave more than one reason)

The 'openly egocentric' category was assigned to students whose primary reaction involved focusing on their own feelings such as wanting to avoid feeling embarrassed. Forty-nine per cent of students chose anonymity since it meant avoiding the worry, awkwardness or discomfort of having to give (negative) constructive feedback, thus focusing on their own feelings. Respondent 7, for example, reported he/she would have felt: "Not so worried about what they would think of me if I wrote something negative" had feedback been anonymous and Respondent 35 stated anonymity to be preferable: "Because then you don't feel like the person is holding a grudge against you if they don't like your opinion."

'Apparently altruistic' reasons were defined as reasons which (on the surface) appeared to focus on the feelings of the recipients i.e. not wanting peers to be upset, but which could also be construed as a focus on the self, as will be explained shortly. Thirty-three per cent of students fell into this category, for example, Respondent 31 who stated: "[I] thought I may offend some members of the group by writing negative feedback" and Respondent 55 suggested: "You don't want to hurt anyone's feelings." These students seemed unaware that anonymity

would not make any difference to the hurt they envisaged their peers would feel.

Validity reasons were defined as expressions of the belief that anonymous comments are more likely to be honest. Twenty-two students (40%) gave reasons of this nature, for example Respondent 10 who enthused: "Then you can give a really in-depth analysis" and Respondent 14 who claimed: "So I could be honest ... also, so people can be honest about me" and Respondent 79 offered: "I think more people would tend to be honest and you would get genuine feedback as a result."

Just under half the students (49%; n = 53) reported they would *not* have preferred to give feedback anonymously, i.e. they preferred the present system of giving face-to-face feedback. Although five students did not give reasons, the following four themes (illustrated in Figure 9.5) emerged:

- **Position of strength**: taking responsibility for one's own comments and a belief that one should be able to handle constructive criticism (n = 28).
- **Pragmatic**: to permit clarification from the person providing the feedback, if necessary (n = 9).
- **Position of non-criticality**: these students were happy to attach their names to feedback because it did not contain anything negative (n = 7).
- **Curiosity**: about knowing who wrote what about them (n = 4).

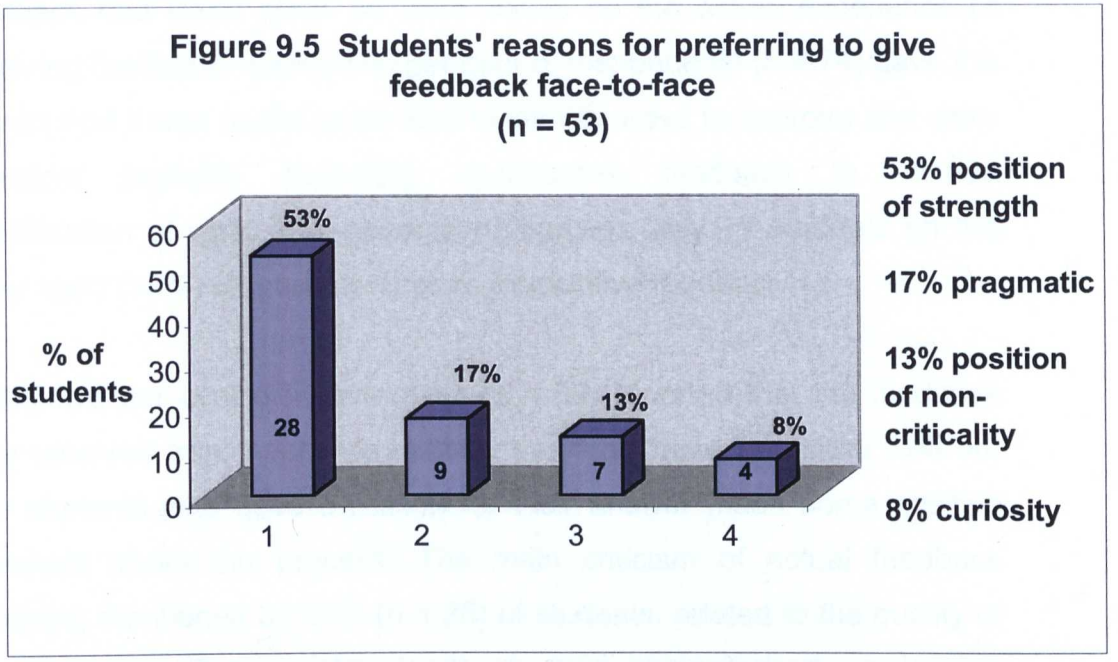


Figure 9.5 Students reasons for preferring to give feedback face-to-face. (N.B. not all students gave reasons for their answer)

I was surprised that so many students preferred the current face-to-face method of providing feedback. However, both Krause and Popovich (1996) and Topping (1998) claim that peer assessment can instil a greater sense of responsibility and motivation in students involved in the process so maybe the process of being involved in setting the criteria made students more willing to take responsibility for providing feedback based on these criteria. This was evident in the comments of 28 students whose mature attitude reflected what I have referred to as a 'position of strength.' For example Respondent 17 commented, "Everyone should be responsible for what he/she says," Respondent 37 who stated: "I feel you should be able to say what you think to a person's face," and Respondent 5 who reported: "I think we are mature enough not to be too negative or cover our true feelings." Another possible interpretation is that students saw the value of the process having experienced it first hand.

Ninety-seven respondents (90%) felt that it was a good idea to receive feedback on their group work skills from peers with whom they worked. It is important to note that students' responses were elicited *after* the

feedback had taken place so were based on the actual experience of receiving feedback. Eighty-two per cent of respondents (n = 74) gave the reason that it was useful to be able to identify ways to improve and were therefore implicitly expecting constructive feedback to include identification of what they needed to improve. Only 17 students on the other hand (18%) expressed a desire for positive feedback.

Eighty-five per cent of respondents (n = 92) reported that the feedback they received from peers was either "very" or "quite beneficial". All but four students who gave a reason for their answer made some positive comment about the process. The main criticism of actual feedback received, mentioned by 24% (n = 26) of students, related to the quality of the feedback. Poor quality feedback was characterised as lacking information about negative aspects of group work behaviour which may have been due to three reasons: firstly, embarrassment about saying negative things to peers; or secondly, an inability to articulate constructive feedback; or thirdly, lack of awareness of the value of including negative elements when providing constructive feedback.

It is unlikely that students were unaware of the importance of constructive feedback given the large number of students (82%) specifying the reason for wanting feedback as being in order to *identify weaknesses*. It could therefore be interpreted that the 24% of students (n = 26) who did not receive constructive feedback did not receive it either because their peers were too embarrassed to be honest in their criticism or did not have the skills to articulate their criticism.

The value of receiving peer feedback clearly depends to a large extent on the quality of the feedback given and, ironically, despite the fact that students wanted information which would allow them to identify how they could develop, in 26 cases (24%) students were *unwilling* or *unable* to give this information to others. If students were to receive more practice in providing formative feedback to increase their skills and confidence, they would have the potential to be an even more valuable resource in

terms of helping to improve each other's group working skills. However, it appears that for some people this feedback needs to be anonymous to encourage the articulation of honest criticism.

Students were also asked how they felt about the process of giving each other marks, i.e. summative peer assessment. Ninety-eight students responded to this question. Eighty-two per cent of students (n = 80) who answered this question reported positive feelings about summative peer assessment. Seventy-one students (72%) felt it was an effective summative tool and students reported that it was a valid method of assessing individual contribution towards the group assignments, for example Respondent 1: "felt this was good; we were able to assess [peers] in a similar way that a tutor would" and Respondent 8 thought: "it was a good way of judging how well a member of the group did."

In particular, they appreciated that students were in a unique position to assess peers' group work skills, for example, Respondent 60 thought it was "a fair system – we are the only ones present in group meetings therefore we are the only ones who can judge" and Respondent 96 felt it was "a good idea. Group members are the ones who understand you the best."

Eighteen students (18%) reported feeling more comfortable or finding it easier to allocate marks than provide feedback, for example Respondent 9 felt it was "not so bad because it is easier to rate with numbers and not words" and Respondent 31 "found this much easier to do than writing feedback for members of the group."

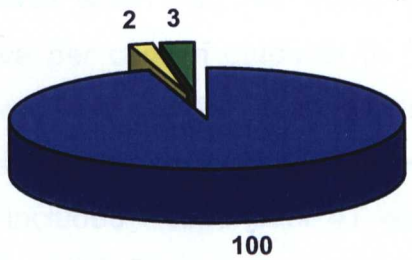
However nine students (9%) expressed some concern about students carrying out assessment, for example Respondent 15 didn't think s/he was "in a very objective position to do so", Respondent 30 preferred "for a tutor to give me marks/grades, not peers" and Respondent 56 felt s/he "did not have the right to be judging my team members." This accords with Orsmond and Merry's (1996) finding that students felt unqualified to

carry out peer assessment. However, it is important to note that only 9% of students expressed this concern; the vast majority of students in the current study did not mention this issue.

Two students did not feel summative peer assessment was necessary. Respondent 3: "thought the written [formative] assessment was more useful" and Respondent 7 felt that it was "not necessary when we had done the first" [i.e. the formative assessment]. It is interesting to note that only two students in the whole of the research over four years considered the option of not assessing group work summatively.

Summative marks were confidential. Students were asked whether or not they thought that the fact that the marks they gave were anonymous was a good idea. One hundred of the 105 students who answered this question said yes (95%). Figure 9.6 shows the number of students who gave each response.

Figure 9.6 Students' preference for anonymous summative assessment (n = 105)



100 students (95%) preferred anonymous summative assessment

2 students preferred summative assessment to not be anonymous

3 did not mind or did not know

Figure 9.6 Students' preference for summative peer assessment to remain anonymous

Of the 100 students who said they felt summative marks should be kept confidential, twelve did not give a reason why. Students' responses fell into the same three categories that emerged when students gave reasons why formative assessment should be kept confidential (see Figure 9.4) and are illustrated in Figure 9.7.

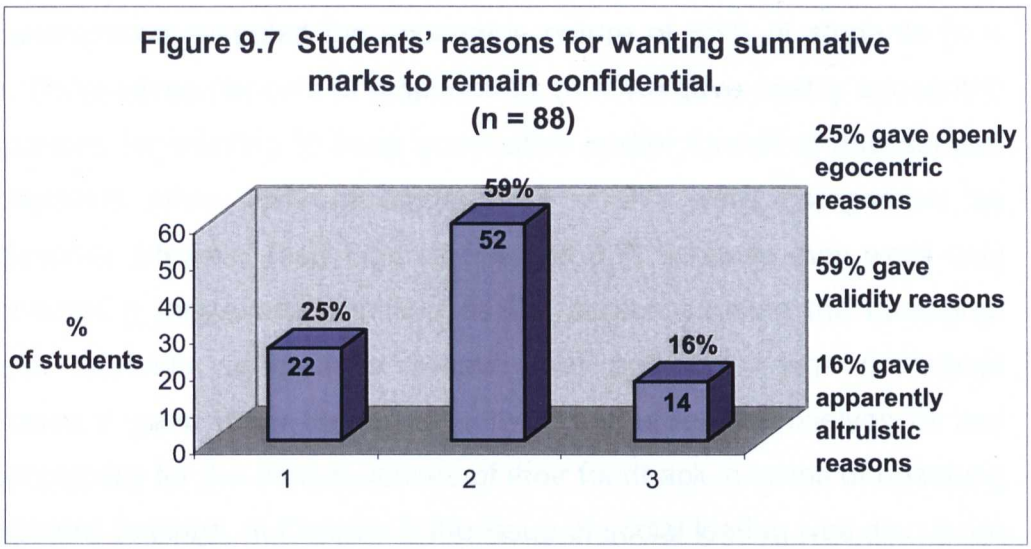


Figure 9.7 Students' reasons for wanting summative marks to remain confidential

Fifty-nine per cent of students (n = 52) preferred summative marks to remain confidential because they believed this ensured greater validity. Students used the words 'honest', 'truthful' and more 'critical' a total of 39 times, for example: "it was easier to be honest" (Respondent 1); "no one knows; it makes it easier to mark truthfully" (Respondent 43); "it is easier to be critical when you are doing it anonymously" (Respondent 63). Twenty-five per cent of students (n = 22) were 'openly egocentric' i.e. they were frank about not wanting to experience any personal repercussions, so preferred to hide behind a cloak of anonymity. These students included Respondent 41 who stated, "You don't want to lose friends over it." Fourteen students (16%) gave 'apparently altruistic' reasons for preferring confidentiality, i.e. they *appeared* to focus on the feelings of others. A common verb used was that they did not want to 'offend' anyone (nine respondents) or similar, for example, "I don't want to dishearten other members of the group whom I didn't give as high marks to" (Respondent 45). Again these students did not seem to realise that their feedback would have the same impact on recipients regardless of whether or not it was anonymous.

A cross-tabulation of students who preferred summative marks to remain confidential and those who would have preferred formative assessment to be anonymous revealed the egocentric nature of 55% of students (n = 59). Thirty-seven per cent of respondents (n = 40) gave openly egocentric responses for wanting to keep summative and/or formative assessment anonymous while 25% of students (n = 27) were categorized as 'apparently altruistic' (see Figures 9.4 and 9.7) because they were only concerned if they were identified as the person causing the upsetting. These students seemed to believe that anonymity was preferable because it gave them the opportunity to unburden themselves of any responsibility for the consequences of their feedback in terms of upsetting recipients' feelings. In Chapter 5 the issue of social loafing was discussed and it was discovered that if individual input was not *identifiable*, then social loafing was more likely to occur since students did not feel accountable for their actions. It also appears that 55% of students would have liked to have been absolved from being accountable for their feedback, preferring to "hide in the crowd".

Although some students made both 'openly egocentric' and 'apparently altruistic' comments, it was the latter categorisation that I found particularly interesting in terms of the lack of self-awareness it reflected. Cook-Greuter (2002, 2004) identifies nine stages of cognitive development, which reflect the meaning-making capacity of people. Her framework is a creative synthesis of previous ego developmental models such as those of Loevinger, 1966; Kohlberg, 1969; and, more recently, Tolbert, 2004. It also reflects some elements of Perry's (1999) scheme of intellectual development. Some of the assumptions shared by such development theories are outlined by Cook-Greuter (2004, p.4) below:

- The unfolding of human potential towards deeper understanding and effectiveness in the world.
- Growth happens through a logical sequences of stages or expanding world views.

- World views evolve from simple to complex and from egocentric to socio-centric to world-centric.
- Later stages can only be reached by transcending the earlier stages.
- People's stage of development influences what they notice and therefore what they can articulate.
- As development unfolds, autonomy and tolerance for difference and ambiguity as well as flexibility, reflection and skills in interacting increase.

According to Cook-Greuter (2002) it is only at Stage 4 that individuals begin to appreciate the consequences of their actions and it is not until later stages that self-awareness develops. Although it would take further research to ascertain students' levels of ego-development, it is possible that 'apparently altruistic' students are at a lower stage of cognitive development. The implications of this are discussed in the final chapter.

Section 3: To what extent does peer assessment of contribution alleviate social loafing?

This section explores the extent to which the peer assessment intervention alleviated social loafing. Students in the sample discussed in Section 2 were involved in two group tasks, producing a poster and making a group presentation. In order for students to assess each other's contribution it was important that they considered their input identifiable to other members of the group (see Latané et al, 1979; Williams et al, 1981; Harkins & Szymanski, 1989; Slavin, 1989b; Karau & Kipling, 1993; Karau & Williams, 1995; Gagne & Zuckerman, 1999; Gillies & Ashman, 2003; Johnson & Johnson, 2003; Price & Harrison 2006).

Ninety-six per cent of students (n = 104) commenting on the poster and 99% (n = 107) of students commenting on the presentation reported that they considered their input easily identifiable to other members of the

group when working on the poster and presentation respectively. This meant that four students did not consider their input identifiable when working on the poster and one student did not consider his/her input identifiable when working on the group presentation. All four students commenting on the poster were working in groups of five which may reflect earlier findings that identifiability becomes more difficult, the larger the group and justified my preference for groups involving no more than four members (see also Karau & Kipling, 1993; Johnson & Johnson, 2003; Kerr & Bruun, 1993; Karau & Williams, 1995). The one student who reported her input was not identifiable in the presentation, missed the presentation due to illness and presented on her own at a later date.

Students were asked to what extent peer assessment had affected the way they worked on the module. One hundred and four students answered this question. Eighty students (77%) said that peer assessment increased their accountability to other group members, which encouraged them to ensure they pulled their weight. For example Respondent 6 "made more of an effort", Respondent 12 reported, "It caused me to work harder within the group, "Respondent 50 "definitely [felt] that others in the group were bearing that in mind which is why we all worked so well." Peer assessment motivated Respondent 62 "to contribute more towards group assignments," and Respondent 63 admitted, "It actually did change my way of working. I started to behave in a certain way so I could contribute as much as I could for the group." Respondent 75 stated, "There was extra pressure on working hard outside lectures which is good!" Respondent 81 stated, "[It] made sure I did equal work and was co-operative," while Respondent 99 felt "this encouraged everyone to work together as a group and not slack off."

Summative peer assessment identified three social loafers so peer assessment did not completely eradicate social loafing. (Social loafers were defined as students whose average peer mark indicated a decrease in the mark awarded for the product of more than 10%.) It did, however, identify the social loafers and penalise them accordingly. Also, since 77%

of students (n = 80) claimed that peer assessment encouraged them to contribute more, it clearly had a positive effect. However, the effects of peer assessment on social loafing could only have been measured accurately if a control group had been incorporated into the design of the study, which undertook the same group tasks without peer assessment. I felt that this was ethically unacceptable since, having asked students about the problems they encountered with group work and identified a key issue as social loafing, I felt I needed to address it to improve the learning experience for all students rather than deny the opportunity to some students. This left me in the awkward situation of having nothing with which to compare the results of my intervention. The only course of action open to me was to replicate questions used in another study that had not introduced a peer assessment mechanism and compare the findings with mine.

Bourner et al (2001) report on the experiences of 73 first year undergraduates undertaking work assessed on the basis of a group project at the University of Brighton. One of the questions they posed was "What positive and negative experiences do first year students encounter when undertaking group project work?" I replicated two questions related to the extent students would like to work in the same group again and the extent to which students worked well together.

Table 9.1 shows students' responses to the question: To what extent would you like to work in the same group on another project?

Response	Bourner et al's results	My results
Very much	13%	55%
A fair amount	30%	35%
A little	20%	10%
Not at all	27%	-

Table 9.1 Comparison of my findings with the findings of Bourner et al (2001)

Forty-three per cent of Bourner et al's students reported they would like to work with the same group members on another project 'very much' or 'a

fair amount' compared to 90% of students in my study. Twenty-seven per cent of Bourner et al's students reported that they would 'not at all' like to work with the same students again compared to none of the students in my study. Table 9.2 compares the extent to which students worked well together in both studies.

The group worked together:	Bourner et al's results	My results
Very well	7%	64%
Well	36%	31%
Satisfactorily	25%	4%
Not too well	23%	1%
Poorly	9%	-

Table 9.2 Comparison of my findings with the findings of Bourner et al (2001)

Forty-three per cent of students in Bourner et al's study reported that they worked 'well' or 'very well' as a group compared to 95% of the students in my study. Thirty-two per cent of students in the Bourner et al study worked 'not too well' or worked 'poorly' as a group compared with 1% of my students. There were clearly major problems associated with working in a group in the Bourner et al study which were identified by the researchers:

> most of the explanations given were associated with group members not making a fair contribution (2001, p.125).

When students in the Bourner et al study were asked how the design or implementation of the project could be improved, two key issues emerged: the need for smaller groups (they found 6/7 too many) and "offering more support/options to those groups where not all the group members are pulling their weight" (2001, p.29). They identified the "issue of passengers" (p.24) as the most problematic aspect of the group work they experienced. Although it is possible that differences in the content and design of the group work account for the discrepancies in students' perceptions between the two studies, most of the explanations given for problems with group work were associated with group members "not

making a fair contribution" in the Bourner et al study (2001, p.125) whereas module evaluation forms used in my study revealed few problems with group work. When students in my study were asked to comment on what they enjoyed and did not enjoy about the module, 80% mentioned group work as one of the three things they most enjoyed and there were no problems related to social loafing recorded. It could, therefore be argued that the introduction of summative peer assessment created conditions in which social loafing is less prominent.

Section 4: Summary and discussion

Having implemented a summative and formative peer assessment intervention, its effectiveness was evaluated by the students involved since I was primarily interested in *students' perceptions.* Eighty-seven per cent of students agreed that being involved in the process of deciding the criteria on which to assess each other's contributions and group skills was a positive thing. Eighty-five per cent reported that formative peer assessment was beneficial, and 67% expressed positive feelings about the experience although a third of students found it difficult to provide criticism either because they could not find anything negative to say, were too embarrassed to criticise peers face-to face or felt they could not articulate their criticism clearly. Twenty-six per cent of students exhibited a strength-based attitude to giving and receiving face-to-face feedback which was characterised as taking responsibility for one's own comments and a belief that one should be able to handle constructive criticism. Although a surprising number of students (49%) chose to maintain the status quo (i.e. provide face-to-face feedback), a significant number of students (51%) would have preferred to give feedback anonymously. The main reason for the latters' choice was that anonymity would allow them to be more honest with their criticisms and avoid what they saw as the awkwardness of criticising their peers face-to-face.

Eighty-two per cent of students expressed positive feelings about the summative peer assessment process in which their marks were awarded anonymously. They acknowledged the effectiveness of summative peer assessment as a tool for assessing group work skills and contribution, which could not be assessed by the tutor and 95% of students felt that summative assessment should remain anonymous to ensure students' gave honest appraisals of their peers' contributions. Consequently, this research goes some way towards creating a pragmatic and evidence-based approach to group assessment in undergraduate courses.

However, although this study found that students reacted positively to group work in which summative peer assessment of contribution to group effort went some way to alleviate social loafing, there are clearly other variables that need to be taken into account. Firstly, the first year students involved in evaluating their experience had all signed up to a module that focused on human communication. Students studying other subjects might not see the benefits of group work when the focus is not on communication and second and third year students who experience too much group work may become bored with it.

Secondly, Thorley and Gregory (1994) suggest that students who are plunged into group work without understanding the group processes involved may flounder. This module introduced students to group work processes and considerable effort was made to "transform a nominal group into a cohesive unit" (Robson, 1994, p.40) in this third cycle of research.

Thirdly, students had the opportunity to practise assessing their peers on two occasions prior to carrying out summative and formative peer assessment of contribution to group effort. They summatively assessed each other groups' posters and presentations and had the opportunity to compare the marks they awarded with the marks I awarded, increasing their confidence as they realised the average difference was 3%. They also had the opportunity to provide formative feedback on posters and

presentations and to compare feedback given. This all takes time, which is justifiable on a module focusing on communication, but other tutors may have other priorities.

Finally, the assessment here was not high-stakes assessment since first year marks do not count towards degree classification. Nevertheless, only 19% of students disagreed or strongly disagreed that peer assessment marks should be a component of a personal final module grade if assessed group work is involved compared with students in a study carried out by Krause & Popovich (1996) in which 47% disagreed or strongly disagreed with this course of action. My findings are more in line with the findings of Stanier (1997) who reported that 74% of the students in his study felt that peer assessment should be an element in modules involving group work.

Chapter 10: Conclusion

This action research project recorded a four-year study involving 794 first, second and third year undergraduate students. It focussed on *students' perceptions* of group work and its assessment. Since much of the existing literature about group work in higher education is presented from a tutor's, employer's or policy maker's perspective and very little from the students' perspective, this study aimed to address this gap.

This chapter is divided into five sections. Section 1 draws together the main findings using seven overarching themes that emerged as a result of analysis on the implications of this study. Section 2 reflects on the espoused theories that underpinned my practice and compares them to my theories-in-use, examining the dilemma this created. Section 3 reflects on the research process using Cranton's (2006) levels of engagement to elucidate the multiple roles I adopted as researcher, co-researcher and teacher/assessor. Section 4 outlines the strengths and limitations of the research. The chapter concludes with Section 5 which discusses further questions that emerge from the study.

Section 1: Significant overarching themes

A number of overarching themes emerged which have implications for policy and practice. The first three themes relate to the initial research questions: What are students' perceptions of group work in higher education; what do they like about it; and what problems do they report? The final four themes relate to the second and third cycles of action research, which focused on peer assessment issues:

i. Group work as a social activity.
ii. Students' need for training in group work.
iii. Students' dislike of social loafing.
iv. Criterion-referenced versus holistic assessment.

v. Students' involvement in setting assessment criteria.

vi. Assessment tensions.

vii. Formative peer assessment.

i. Group work as a social activity

This study suggests that group work could be viewed as a social activity that can foster the development of a community of practice. In the first cycle of research students identified two key learning benefits: learning how to work with others and learning from others through discussion. They also appreciated the support they received from other members of the group. They reported that they enjoyed the social aspect of group work, which made learning more fun. Group work encouraged a sense of community that students felt was particularly important at the beginning of their first year of university. This perceived sense of community was of particular interest in that it seems to set a foundation on which to construct what has been referred to as a cultivated community of assessment practice (O'Donovan et al, 2008).

O'Donovan et al (2008, p.205) propose a "nested hierarchy" of four models which co-exist in higher education and which map approaches to developing students' understanding of assessment standards. Their 'laissez-faire' model reflects the traditional practice in which students gradually absorb tacit standards serendipitously, as a result of discussion with tutors. Their 'explicit' model is based on the assumption that careful articulation is sufficient to clarify assessment standards, which are then passively presented to students. The authors suggest, however, that the laissez-faire model is unlikely to work due to the erosion of informal opportunities for discussion with tutors as a result of the massification of higher education and the resultant increase in student:staff ratio. They also suggest that explicit articulation is not enough since social processes play a key role in establishing standards and in the development of students' understanding of assessment and feedback. As a result, they suggest a 'social constructivist' model involving students engaging in the

process of assessment and a 'cultivated community of practice' model in which tacit standards are communicated to students through the latters' "participation in informal knowledge exchange networks seeded by specific activities." O'Donovan et al (2008) note, however, that there is scant evidence in the literature as to what these "seeded activities" might entail.

From my study it could be argued that introducing formative and summative peer assessment of contribution to group work could be a way of fostering such knowledge exchange networks. It is the ideal assessment activity because it can *only* be carried out by students themselves.

However, students in this study reported difficulties with arranging meetings due to conflicting timetables, work patterns and (specific to the university in which the research was carried out), travelling between campuses. Students need time and space to meet when carrying out group work so resources need to be allocated and planning needs to take place either to timetable group work sessions alongside taught sessions or to encourage students to engage in group work outside taught sessions. These findings suggest that policies on time and space management may merit revision in order to facilitate group work in university settings.

ii. Students' need for relevant training

This study found that students felt they lacked knowledge about how to work effectively as groups, thus suggesting a need for training. Students on a module on Human Communication who took part in the first cycle of research were not confident that they had adequate skills to manage group work effectively. This supports Newman and Nelder's (1994) claim that students "must have opportunities to reflect on the performance of the process as well as [group work] outcomes" (p.48). As a result, much more time in subsequent runs of the module was dedicated to helping students understand how groups function. It is suggested that if tutors opt

to use assessed group work, they need to be prepared to devote considerable time to developing students' group processing skills. This research also found that students generally viewed conflict as negative, so training needs to involve discussion about the benefits of constructive confrontation and include conflict resolution processes.

Also, although 85% of the students acknowledged the current feedback process to be 'very' or 'quite beneficial', clearly "an important determinant of the effectiveness of formative assessment is the quality of the feedback received by learners" (Yorke, 2003, p.482). It would appear that workshops to improve students' skills and confidence in providing constructive feedback are important in the light of these findings.

iii. Students' dislike of social loafing.

In this study the expression 'social loafing' has been introduced to describe the behaviour of students who are considered by their peers as not 'pulling their weight' in group work. Students used a number of terms to describe this phenomenon, but since 'social loafing' is used frequently in the literature to describe the reduction in effort of people working collectively (Latané et al, 1979), it was adopted here as an umbrella term to describe this trend.

It was apparent that the most common complaint about group work in this study is peers who do not pull their weight. Students' concerns focused on the issue of fairness; when group members are all given the same group mark, students see this as unfair because it fails to discriminate between the different levels of contribution group members make. When given the choice between receiving one group mark or using summative peer assessment to arrive at individual marks based on contribution, students chose the latter. These findings suggest that it is important for tutors to take steps to reduce social loafing when group work is to be assessed so that students feel that assessment is fair. Ensuring groups are small and that identifiability is established through summative peer

assessment of contribution improves students' perceptions of the fairness of assessed group work. It is also suggested that individual marks be weighted around the product mark and that an attempt be made to establish inter-group competition to encourage cohesion between in-group members.

iv. Criterion-referenced versus holistic assessment

The results of this study show that students prefer criterion-based assessment to holistic peer assessment. Ninety-three per cent of respondents felt the former was less 'personal', more 'objective' and facilitated the process of differentiating individuals' skills and contributions. These findings support the use of the 'explicit model' in O'Donovan et al's (2008) nested hierarchy of approaches to assessment since the articulation of criteria clearly improved students' understanding of assessment. Students also felt that criterion-based assessment opened up the opportunity to receive feedback on each criterion included.

v. Students' involvement in setting assessment criteria

The use of criterion-referenced peer assessment was combined with a social constructivist approach; the students involved generated their own criteria. This was found to be a key element in developing students' understanding of the assessment process and emphasises the importance of the nested approach to assessment proposed by O'Donovan et al (2008).

vi. Assessment tensions

The assessment of group work produced two main tensions: firstly, tension between fairness, transparency and validity of peer assessment of contribution to group work; and secondly, tension between the

delegation of control over the assessment process and the tutor's role as assessor.

The Assessment Standards Knowledge Exchange's (ASKe) manifesto for change states: "learners who engage in assessment are entitled to fair and transparent standards" (Price et al, 2008). It could be argued that firstly, tutors who currently award one group mark for assessed group work are not conducting assessment fairly, since not all students contribute equally to group work, and secondly, those who do not include students in the discussion of criteria for assessment are not conducting assessment transparently.

However, although students in this study did engage in the assessment process and appreciated the fairness and transparency it afforded, questions over validity remain. Summative peer assessment by students clearly only makes sense if students are capable of assessing at least as validly and reliably as tutors. On the one hand, there is considerable evidence in the literature to support the belief that they are (see for example: Burnett & Cavaye, 1989; Orpen, 1982; Topping, 1998; Oldfield & Macalpine, 1995; Norton, 1992; Falchikov, 1986, 1995; Hughes and Large, 1993; Stefani, 1994; Butcher and Stefani, 1995; Scoffield & Brindley, 1998; Magin, 2001; Langan et al, 2005; McAuley 2009b). On the other hand, there is also evidence that students are not reliable and valid markers (for example, Mowl & Pain, 1995; Maguire and Edmondson, 2001). Also, there is little discussion of *intra*-group assessment of contribution such as is the focus of my research. Assessing other students' posters and presentations in a relatively impersonal large class setting is very different to assessing peers' contribution to a group assignment in a small group setting in which relationships have developed over time. This study found that small groups whose members work together over a period of time develop relationships that group members perceive may interfere with the ability to award marks objectively and identified students' concerns about biased marking as a consequence.

It would seem from these findings that *summatively* assessed group work should not comprise a major part of high stakes assessment. Although I continue to use group work in some part of each module I teach, I do not use *assessed* group work to account for more than 25% of a module mark. Findings also suggest that when summative assessment is used, confidentiality of marks is important: ninety-five per cent of students in this study preferred confidentiality with 59% stating that they gave more honest marks because their identities were protected.

A second assessment tension exists between the delegation of control over the assessment process and the tutor's role as assessor. Students in this study expressed appreciation at being given more control over the assessment process. This resulted in a shift in power from tutor to student control which is also reflected in Klenowski's (1995) study; she discusses how negotiation of assessment criteria had an impact on the student-teacher relationship in her study resulting in a feeling of a redistribution of *power* (my italics). This shift in power can threaten the traditional hierarchical nature of the tutor-student relationship which some may prefer to maintain.

McDowell & Harman (2008) discuss the multiple assessment discourses that shape the identities of assessors in higher education. Their analysis suggests assessors construct four positions between which tutors shift: 'traditional teacher', 'expert guide', 'professional assessor' and 'gate-keeper.' Involving students in the assessment process introduces a range of voices traditionally excluded from the community of assessment practice which may frame an alternative discourse and which may challenge the identities of assessors in these four positions or roles. There is already evidence that tutors who feel comfortable with the traditional tutor-student hierarchical relationship feel threatened by the autonomy delegated to students in terms of assessment and disenfranchised by the threat to their identity or to what they see as their legitimate role as assessors (for example, Falchikov, 2005; Ecclestone,

2009). Some tutors may reject the process described in this study as a consequence.

vii. Formative peer assessment

This study underlines the importance of the role of formative peer assessment in assessment *as* learning but raises questions as to whether or not peer feedback should be provided on a confidential basis.

The rationale for introducing face-to-face formative assessment in this study reflected my belief that it would instil a sense of responsibility in students. I felt that not only should students be identifiable as the source of, and therefore responsible for the feedback provided, but that it would make them think more carefully about how they worded any criticism. Providing constructive feedback is an important skill and particularly relevant to the world of work (see for example Topping, 1998; Gillies & Ashman, 2003). Moreover, Sadler (1989) suggests: "providing direct and authentic evaluative experience is a necessary condition for the development of evaluative expertise and therefore for intelligent self-monitoring" (p.143). Findings from this study suggest that students need to learn how to voice their ideas sensitively, constructively and clearly, how to defend their point of view (see also Bloxham & Boyd, 2007) using concrete evidence, and how to deal with any awkwardness that may arise as a result. It is a form of assessment *as* learning (Earl, 2003) since students learn from the process of carrying out assessment itself. Some students felt somewhat daunted when faced with the challenge of having to think about how to articulate and justify their praise and criticism, and this study found that this challenge is magnified when feedback is not undertaken on a confidential basis.

Also, since it is important that feedback is valid, if students avoid giving criticism because of the repercussions they envisage, then the quality of the feedback is compromised. Since the majority of students (82%) who said it was a good idea to receive peer feedback in this study, wanted to

be able to identify weaknesses rather than strengths from the process, it would potentially be better if the process were anonymous to encourage those constructing feedback to be honest in their criticisms. The number of students in this study who reported that their feedback would have been more honest had it been anonymous reinforces the need for anonymity.

When the above findings and supporting evidence from the literature are taken into account, it could be argued that the process needs to be introduced early because it gives group members a chance to consider their own contribution and see how they could improve (see also Dominick et al 1997; Druskat & Wolff, 1999; Reilly et al, 1996; Smither et al, 1995; Brook & Ammons, 2003). Also, if introduced early, it is more likely to become part of a student's repertoire of tools to aid learning.

In addition, given the question mark over marker bias emanating from peer reciprocation identified above, it could be argued that *formative* peer assessment is "a safer vehicle to invite learners to take a more central position in assessment praxis" (O'Donovan et al 2008).

Section 2: Reflection on espoused theories and theories-in-use

It was argued in Chapter 3 that an action research approach was used because it is a learning vehicle in itself - an intrinsic, integral part of professional practice (Winter, 1989). This research has therefore given me the opportunity to study group work and peer assessment on both theoretical and practice-based levels thus contributing to my professional development.

The literature reviews and research carried out to establish a greater understanding of group work and peer assessment allowed me to consider the theories that underpin my practice. Before they were undertaken, my rationale for using group work was based on two sorts of

implicit theories. The first involved what Argyris and Schön (1974, p.6) refer to as "espoused theories" which refer to theories that I have read about in the past, incorporated into my way of thinking which I truly believe to govern my actions and yet whose origins are now lost. The second sort derive from what Argyris and Schön (1974, p.6) refer to as implicit "theories-in-use" which reveal what I really believe through my practice and which may contradict my espoused theories. It was important to me to reforge my links with espoused theories and to reflect on my theories-in use in order to articulate and interrogate the latter.

As a result of reflection I was able to rekindle the link between social constructivism as espoused theory and recognise that it underpinned my predilection for group work as a theory-in-use. I also discovered that my social constructivist epistemology resonated with the pragmatic methodological approach outlined by Morgan (2007) since both stress the view of knowledge as a dialectical process of meaning-making. The latter ensured a workable line of action (implementing peer assessment) was based on warranted assertions (feedback from students).

On a personal level, carrying out the research presented me with a dilemma relating to the nature of assessment, the extent to which it drives learning and the consequences for promoting autonomous learning. This forced me to question my role as an assessor and the assessment practice of the institution in which I work.

My observations and experience of working with undergraduate students suggest that assessment drives learning for the majority. This espoused theory is confirmed by a number of studies which suggest that "students' behaviour and attitude to learning is highly influenced by the assessment system" (Freeman, 1995, p.290) and that assessment is the tail that wags the dog (Gibbs, 1992). (See for example Webster et al, 2000; Brown et al, 1997; Bloxham & Boyd, 2007; Ramsden, 1992; Ravenscroft et al, 1999.)

Consequently, I believe that if the product is not assessed at all, most students will lack the incentive to take group work seriously in the first place. Moreover, when only the end product is assessed in group work, it seems that students will largely ignore the learning process and focus only on the learning outcome (see also Gillies & Ashman, 2003 and McWhaw et al, 2003). I therefore agree with Strachan & Wilcox's (1996, p.348) assertion, "process must be stressed as equally important to product, for it is in the process that the actual learning takes place."

Seventy-seven per cent of the students in the current study reported that peer assessment made them "pull their weight" and 86% in a parallel study reported that the fact that they were tested on set reading on a weekly basis made them work harder (McAuley, 2009b). Fifty-two per cent of students said they would probably or definitely not have read weekly set readings had they not been assessed, while 25% were not sure if they would have read them. So, there is considerable evidence that undergraduate students are extrinsically motivated by summative assessment.

However, the fact that students are extrinsically motivated left me with a discordant note sounding in my ears. This study makes it clear that fostering student autonomy is important to me and I have gone to great lengths to delegate responsibility to students in terms of choices of what to do to alleviate social loafing and what assessment criteria to use. And yet my approach to assessment, described in the previous paragraph, does not sound like a tutor trying to promote student autonomy. This realisation forced me to review the notions of autonomy embedded within my assessment practice.

The conflicting epistemologies that underpin assessment practice have been highlighted by Ecclestone (2002). She claims that practitioners' views of student autonomy and motivation often combine 'espoused theories' with implicit 'theories-in-use' which she describes as "underlying but *unrecognised* theories" revealed through practice (Ecclestone, 2002,

p.28, my italics). It appeared initially that my 'espoused theories' might be in conflict with my 'theories-in-use', which would have explained my uncomfortable feeling of disequilibrium. While espousing a constructivist theory of learning that assumes students are intrinsically motivated, my practice seemed to reflect a behaviourist notion of assessment: students are viewed as extrinsic learners whose learning is driven and shaped by summative assessment.

However, closer examination revealed two things. Firstly, it appeared that my feeling of unease was based on the fact that I was approaching assessment as a constructivist who is also a pragmatist responding pragmatically to the reality that the majority of students are extrinsically motivated. Secondly, the formative and summative assessment process described in this study is grounded in my constructivist epistemology which can be explained using Ecclestone's (2002) framework of autonomy.

The formative feedback mechanism in which students were involved encouraged development through two levels of autonomy which act as a form of scaffolding: it encouraged procedural autonomy through the negotiation of assessment criteria and the acquisition of confidence in the language of assessment and it encouraged personal autonomy through providing conditions for positive interdependence between group members. This developed students' ability to appraise their skills and weaknesses. Students did not describe the experience in terms of a hurdle to be addressed (extrinsic motivation) but instead exhibited intrinsic motivation by describing formative assessment as an opportunity to learn and develop. Perhaps the fact that I facilitated this apparent shift in their attitudes reflects a rather different role for McDowell & Harman's (2008) 'expert guide' to which I aspired, one whose scaffolding gradually withdraws to be replaced by scaffolding provided by other students in group contexts.

Section 3: Critical reflection on the research process

Because of the nature of this study, it is important to reflect critically on the impact of my multiple roles in relation to design, participants and outcomes. This is enhanced by thinking in terms of Cranton's (2006) three levels of reflective engagement as premise driven researcher, process driven action researcher and content driven teacher/assessor. Firstly, my role as researcher demanded a more detached approach that could have been impeded by my choice of an action research strategy which requires the adoption of a more participative role. On the one hand, it could be argued that my increased level of engagement in the study was more likely to result in a set of shared meanings, or intersubjectivity, in terms of interpreting students' perceptions of group work. This was my explicit intention. However, it could be argued that since an objective separateness was less likely to be maintained between me in my role as action researcher and students as participants, findings may reflect my values to a larger extent than would have been the case had a more detached researcher carried out the study.

Moreover, the complexity of two interwoven levels of engagement (I was simultaneously trying to remain objective while arriving at an intersubjective understanding of students' perceptions of group work) was compounded by a third level – that of teacher/assessor. Students may have been unable to separate my role of researcher from role of assessor despite assurances that assignments would be marked before questionnaires (which were anonymous) were analysed. Also my role as teacher/assessor impacted on the study in terms of how and when I collected data. This is discussed further in the section below.

Section 4: Limitations and strengths

Despite one of my supervisor's assertions that my research should take priority over my teaching, in reality my research had to fit around my

teaching commitments. As a result, some difficult choices had to be made. It would have been advantageous to have followed up questionnaires by interviewing a sample of students in order to gain more in-depth understanding of their views. Had this happened, students would have had the opportunity to clarify anything they did not understand in the questionnaires or expand their answers and I would have been able to follow up intriguing answers on questionnaires that consequently remained unexplored. On the one hand, lack of interviews could be seen as a weakness in this study; on the other hand, questionnaires had the advantage of being anonymous and completed at a distance from me in the classroom so students may have been more likely to respond more truthfully. Students may have been less able to separate my role of researcher from role of assessor had they been interviewed face-to face.

I decided to sacrifice depth (to some extent) for breadth because I aimed to elicit perceptions generally representative of undergraduate students. However, since the evaluation of the formative and summative assessment mechanism put in place was limited to first year students on a Human Communication module, it is questionable as to how generalisable the findings reported in Chapter 9 might be. The next cycle of my research will expand the sample to second and third year undergraduates in order to attempt to make the findings more generalisable to the population.

In chapter 3 I underlined the importance of ensuring inter-rater reliability. which is important within Morgan's (2007) pragmatic approach to carrying out research. I was undertaking research to assess the workability of a potential line of action (in this case the implementation of a peer assessment intervention) based on warranted assertions (my interpretation of students' perceptions of the process). In other words my aim was to use inter-rater reliability as a check to make sure that the understanding of students' perceptions I arrived at was corroborated by other sources and did not derive subconsciously from my value system. However, although various colleagues coded samples of text and

intercoder reliability was high, the categories which colleagues used to code text had been decided by me. Had colleagues been asked to determine the categories themselves, these may have been different. As Dey (1999) and Willig (2001) point out, categories do not simply 'emerge' from the data because they do not exist before the process of categorisation; rather they are suggested by the researcher during the research process.

Further, terms such as 'gut feeling' were adopted without consideration as to the different interpretations possible. When students were asked to choose between criterion-referenced assessment and 'gut-feeling,' perhaps the latter had negative connotations for them whereas the term has positive connotations involving professional holistic judgement to me.

The scope of this thesis has been rather large and the journey for the reader rather circuitous. Having started with an exploration into students' perceptions of group work, it identified social loafing as students' key concern, explored ways to alleviate the problem, presented students with a range of options and ended with an examination of students' perceptions of the formative and summative peer assessment process they chose to implement. In so doing it suggests some steps towards a pragmatic and evidence-based approach to the assessment of group work in undergraduate courses.

In addition, the experience of carrying out the research felt somewhat disjointed as I struggled to collect data on students' perceptions in a range of modules without taking valuable teaching time out of sessions. Decisions on how to conduct the research may have been different had I not also been involved in teaching. Also, research questions could not be outlined in advance due to the nature of the study which did not begin with a set of propositions to test. Lincoln & Guba refer to this kind of research as "emergent" (1985, p. 208). Since I value organisation and clarity of structure and direction in my professional life, it is ironic that this sort of research is the least likely to provide these.

Three issues in particular were very frustrating. Firstly, writing a research proposal was difficult because until the problems students experienced with group work were identified, strategies to address these problems could not be planned. It felt very uneasy starting a process which could not be clearly delineated in advance. Secondly, the mainstream models of action research were frustrating because they did not reflect the messiness of the action research process in real life. Thirdly, instead of being able to undertake one literature review, a number of literature reviews needed to be undertaken along the way in order to answer questions that emerged as the research unfolded. This is a feature of action research methodology which necessitates the juggling of time between data collection and analysis from one action research cycle with further literature searches for the next cycle.

Although I was explicit in stating that the focus of this study was student perceptions, I did not define what I meant by perceptions and rather limitedly focused on likes and dislikes. In addition, although the students' view of assessment of group work is a valuable view, it could be argued that it is a narrow perspective and that assessment decisions taken by students may have been based on personal reasons rather than professional judgement. However, detailed reviews of literature ensured that decisions were also viewed in the light of research findings.

Also, it could be suggested that a further limitation arises from the fact that the student sample kept changing. Although research was repeated with a number of samples of students on occasions, decisions made by previous groups were sometimes imposed on subsequent groups in the assumption that the latter would have made the same decisions as previous groups. This was unavoidable because it was impossible to carry out all the cycles in one academic year due to my teaching commitments; another consequence of my multiple roles. It is not possible to ascertain the impact this had on the study.

In Chapter 3 I underlined the importance of a pragmatic approach. It was important to me that I was able to do something to improve my teaching practice. Although I have introduced a system that seems to minimise social loafing, the summative assessment tool is still at an early stage.

Despite the above limitations and although there are a number of studies that describe methods of using peer assessment of contribution to group effort to arrive at an individual summative mark, to date, no study does all of the following:

- Involves a large sample of students
- Allows the students to choose which method of assessment to use
- Allows the students to devise the criteria for assessing peers' contribution
- Implements a peer assessment method which includes both a summative mark and formative peer feedback
- Evaluates the process both quantitatively and qualitatively from the students' perspective.

Section 5: Further research

This research has uncovered new questions, which would need many more action research cycles to explore and answer. What, for example is the quality of the students' formative feedback provided by the respondents discussed in the study? Although 85% of students in this study found the feedback from peers beneficial, 14% of students reported that feedback was 'not very beneficial' and one student reported that it was 'not at all beneficial'. A follow up study is required to see if ensuring the feedback is anonymous and providing workshops to increase feedback skills improves the quality of the feedback provided.

A second question relates to the perceived difference in ego-developmental levels exhibited by those with a strength-based attitude to

formative feedback in comparison with those I term 'openly egocentric' or 'apparently altruistic.' Is this perceived difference justified?

Thirdly, the term social loafing has a negative connotation and implies a laziness which may not be true of some students. Lack of contribution could be due to cultural differences, lack of confidence, personal problems, lack of English language skills or lack of ability instead. What can be done to differentiate between the indolent and those lacking in confidence or ability and to help such students? Or perhaps social loafing reflects a position of principle adopted by students when the group takes a direction that the "social loafer" is not prepared to take. Although there is some evidence that self-beliefs related to one's feeling of uniqueness are a significant component of social loafing (see Charbonnier et al, 1998 who found that social loafing is, in fact, typical of people who are motivated to assert their individuality and uniqueness), the impact of dispositional factors on social loafing has so far been neglected in the literature.

Fourthly, to what extent is the summative assessment tool produced a valid and reliable instrument? This study reported students' concerns with the possibility of biased marking due to friendship patterns and mutual like/dislike arising from working in a group. Although two studies found reciprocity effects to be negligible (Montgomery, 1986: Magin, 2001) there is a dearth of research into the bias resulting from social interaction and relationships between raters and ratees in group settings.

So many questions still to answer. The journey has, however, left me with a renewed enthusiasm for teaching and a feeling of excitement that there is still much to find out and improve.

References

Abrami, P.C., Chambers, B., Poulsen, C., De Simone, C., D'Apollonia, S. & Howden, J. (1995) *Classroom Connections: Understanding and Using Cooperative Learning*, London: Harcourt Brace

Agyemang, G. & Unerman, J. (1998) Personal skills development and first year undergraduate accounting education: a teaching note, *Accounting Education*, 7, (1) pp.87-92

Anderson, J.B. & Freiberg, H.J. (1995) Using self assessment as a reflective tool to enhance the student teaching experience, *Teacher Education Quarterly,* 22, pp. 77-91

Argyris, C.A. & Schön, D.A. (1974) *Theory into Practice*, London: Jossey-Bass

Bales, R.F. (1970) *Personality and Interpersonal Behaviour*, New York: Holt, Reinhart & Winston

Ballantyne, R., Hughes, K. & Mylonas, A. (2002) Developing procedures for implementing peer assessment in large classes using an action research process, *Assessment and Evaluation in Higher Education*, 27 (5) pp. 427 - 441

Barfield, R.L. (2003) Students' perceptions of and satisfaction with group grades and the group experience in the college classroom, *Assessment and Evaluation in Higher Education,* 28 (4) pp. 355 – 369

Bennett, N. (1991) Cooperative learning in classrooms: Process and outcomes, *Journal of Child Psychology and Psychiatry*, 32, pp. 581 – 594

Biggs, J.B. (2003) Teaching for Quality Learning at University Buckingham: SRHE & Open University Press

Black, P. & Wiliam, D. (1998) Assessment and Classroom Learning, *Assessment in Education*, 5 (1) pp. 7 – 74

Bloxham, S. & Boyd, P. (2007) *Developing Effective Assessment in Higher Education*, Maidenhead: Open University Press

Bloxham, S. & 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) pp. 721- 733

Boud, D. (1989) The role of self-assessment in student grading, *Assessment and Evaluation in Higher Education*, 14 (1) pp. 20 – 30

Boud, D. & Falchikov, N. (eds.) (2007) *Rethinking Assessment in Higher Education,* London: Routledge

Bourner, J., Hughes, M. & Bourner, T. (2001) First-year undergraduate experiences of group project work, *Assessment and Evaluation in Higher Education,* 26 (1)

Brock-Utne, B. What is educational action research? In Kemmis, S. & McTaggart, R, (1988) *The Action Research Reader*, 3rd edn., Victoria: Deakin University Press

Brooks, C. & Ammons, J. (2003). Free Riding in Group Projects and the Effects of Timing, Frequency and Specificity of Criteria in Peer Assessment, *Journal of Education for Business*, 78 (5) pp.268 - 272

Brown, S. (1998) *Peer Assessment in Practice*, SEDA paper 102, Birmingham: SEDA

Brown, G., Bull, J. & Pendlebury, M. (1997) *Assessing Student Learning in Higher Education*, London: Routledge

Brown, S. & Glasner, A. (1999) *Assessment Matters in Higher Education: Choosing and Using Diverse Approaches*, Buckingham: SRHE & Open University Press

Brown, S. and Knight, P. (1994) *Assessing Learners in Higher Education*, London: Kogan Page

Brown, G. & Pendlebury, M. (1992) *Assessing Active Learning* (Module II, Part 1) Sheffield: CVCP

Brown, S., Sambel, K. and McDowell (1998) What do students think about peer assessment? in Brown, S. (ed) (1998) *Peer Assessment in Practice*, Birmingham: SEDA

Bruffee, K.A. (1993) *Collaborative Learning: Higher Education, Interdependence and the Authority of Knowledge*, Baltimore: John Hopkins University Press

Bryan, C. (2004) *Assessing Group Practice*, Birmingham: SEDA

Bryan, C. & Clegg, K. (eds.) (2006) *Innovative Assessment in Higher Education*, London: Routledge

Bryman, A. (1998) *Quantity and Quality in Social Research*, London: Unwin Hyman

Burnett, W. & Cavaye, G. (1980) Peer assessment by fifth year students of surgery, *Assessment in Higher Education*, 5, pp.273 – 278

Burns, R.B. (2000) *Introduction to Research Methods*, 4th edn., London: Sage

Burrell, G. & Morgan, G. (1979) *Sociological Paradigms and Organisational Analysis*, London: Heinemann

Bushell, G. (2006) Moderation of peer assessment in group projects, *Assessment and Evaluation in Higher Education*, 31 (1) pp.91- 108

Butcher, A.C. & Stefani, L.A.J. (1995) Analysis of peer-, self- and staff-assessment in group project work, *Assessment in Education: Principles, Policy and Practice*, 2 (2) pp.165 – 186

Butler, R. (1988) Enhancing and undermining intrinsic motivation: the effects of task-involving and ego-involving evaluation on interest and performance, *British Journal of Educational Psychology*, 58, pp.1 - 14

Carr, W. & Kemmis, S. (1986) *Becoming Critical: Education, Knowledge and Action Research,* Basingstoke: Falmer Press

Charbonnier, E., Huguet, P., Brauer, M. & Monteil, J. (1998) Social loafing and self-beliefs: People's collective effort depends on the extent to which they distinguish themselves as better than others, *Social Behaviour and Personality*, 26 (4) pp.329-340

Cheng, W. & Warren, M. (1997) Having second thoughts: student perceptions before and after a peer assessment exercise, *Studies in Higher Education*, 22 (2)

Cheng, W. & Warren, M. (1999) Peer and Teacher Assessment of the Oral and Written Tasks of a Group Project, *Assessment and Evaluation in Higher Education,* 24 (3)

Cheng, W. & Warren, M. (2000) Making a Difference: using peers to assess individual students' contributions to a group project, *Teaching in Higher Education,* 5 (2) pp. 243-255

Claxton, G. (1995) What kind of learning does self-assessment drive? Developing a "nose" for quality: Comments on Klenowski, *Assessment in Education,* 2 (3) pp. 339 – 343

Coffey, A. & Atkinson, P. (1996) *Making Sense of Qualitative Data: Complementary Research Strategies*, London: Sage

Cohen, L. & Manion, L. (2004) *Research Methods in Education*, 4th edn., London: Routledge

Cohen, L., Manion, L. and Morrison, K. (2007) *Research Methods in Education*, 6th edn., London: Routledge

Colbeck, C.L., Campbell, S.E. & Bjorklund, S.A. (2000) Grouping in the Dark: What college students learn from group projects, *The Journal of Higher Education,* 71 (1) pp.60 - 83

Conway, R., Kember, D., Sivan, A. & Wu, M. (1993) Peer assessment of an individual's contribution to a group project, *Assessment and Evaluation in Higher Education*, 11 pp. 45 – 56

Cook-Greuter, S.R. (2002) A detailed description of the development of nine action logics in the leadership development framework: adapted from ego development theory. Available at http://www.Cook-Greuter.com. Accessed on 30 January 2009

Cook-Greuter, S.R. (2004) *Making the case for a developmental perspective*, Wayland, MA: Industrial and Commercial Training

Cooper, N.J. (2000) Facilitating learning from formative feedback in level 3 assessment, *Assessment and Evaluation in Higher Education*, 25, (3) pp. 279 – 291

Cowan, J. (1998) *On Becoming an Innovative University Teacher,* Buckingham: SRHE & Open University Press

Cowan, J. (1998/2006) *On Becoming an Innovative University Teacher,* 2nd edn., Buckingham: SRHE & Open University Press

Cowie, H. & Rudduck, J. (1990) Learning from one another: The challenge, in H. Foot, M. Morgan, & R. Shute (eds.) *Children helping children* (pp. 235 – 255) Chichester: John Wiley & Sons

Cranton, P. (2006) *Understanding & Promoting Transformative Learning*, San Francisco: Jossey-Bass

Creswell, J.W. (2007) *Qualitative Inquiry and Research Design*, 2nd, edn., London: Sage

Denzin, N. K. & Lincoln, Y.S. (2005) *Handbook of Qualitative Research,* 3rd edn., London: Sage

Dey, I. (1993) *Qualitative Data Analysis: a use friendly guide for social scientists,* London: Routledge

Dillenbourg, P., Baker, M., Blaye, A. & O'Malley, C. (1996) The evolution of research on collaborative learning in Reimann, P. & Spada, H. *Learning in Humans and Machines*, Oxford: Elsevier

Dochy, F.J.R.C. & McDowell, L. (1997) Assessment as a tool for learning, *Studies in Educational Evaluation*, 23, pp. 279 - 298

Dochy, F., Segers, M. & Sluijsmans, D. (1999) The Use of Self-, Peer and Co-assessment in Higher Education: a review, *Studies in Higher Education*, 24, (3)

Doise, W. & Mugny, W. (1984) *The Social Development of the Intellect*, Oxford: Pergamon

Dominick, P.G., Reilly, R.R. & McGourty, J.W. (1997) The effects of peer feedback on team member behaviour, *Group & Organisational Management*, 22 (4) pp. 508 - 520

Druskat, V.U. & Wolff, S.B. (1999) Effects and timing of developmental peer appraisals in self-managing groups, *Journal of Applied Psychology*, 84 (1) pp. 58 - 74

Earl, N.M. (2003) *Assessment as Learning*, Thousand Oaks, CA: Corwin Press

Earl, S.E (1986) Staff and peer assessment – measuring an individual's contribution to group performance, *Assessment and Evaluation in Higher Education,* 11 pp. 60 –69

Ecclestone, K. (2001) I know a 2:1 when I see it: understanding criteria for degree classification in franchised university programmes, *Journal of Further and Higher Education,* 25 (3) pp. 301 – 313

Ecclestone, K. (2002) *Learning Autonomy in Post-16 Education: the politics and practice of formative assessment*, London: Routledge

Ecclestone, K. Personal correspondence, 25 September 2008 & 18 March 2009

Ecclestone, K. & Swann, J. (1999) Litigation and Learning: tensions in improving university lecturers' assessment practice, *Assessment in Education: Principles, Policy and Practice*, 6 (3) pp. 377 – 390

Elliott, J. (1978) What is Action Research in Schools? in Kemmis, S. & McTaggart, R. (1988a) *The Action Research Reader*, 3rd edn., Victoria: Deakin University Press

Elliott, J. (1982) Action-research: a framework for self-evaluation in schools, Working Paper No. 1, *Teacher-Pupil Interaction and the Quality of Learning,* London: Schools Council

Elliott. J. (1991) *Action Research for Educational Change*, Milton Keynes: Open University Press

Elliott, R., Fischer, C.T. & Rennie, D. L. (1999) Evolving guidelines for publication of qualitative research studies in psychology and related fields, *British Journal for Clinical Psychology*, 38, pp.215 - 29

Falchikov, N. (1986) Product comparisons and process benefits of collaborative peer group and self-assessments, *Assessment and Evaluation in Higher Education,* 11, pp. 146 – 166

Falchikov, N. (1993) Group process analysis: self and peer assessment of working together in a group, *Educational and Training Technology*, 30, pp. 275 - 284

Falchikov, N. (1998) Group process analysis in , S. Brown & P. Dove (eds.) *Self and Peer Assessment*, pp. 15-27, Birmingham: Standing Conference on Educational Development, Paper 63

Falchikov, N. (1995) Peer feedback marking: developing peer assessment, *Innovations in Education and Training International*, 32, pp.175-187

Falchikov, N. (2001) *Learning Together*, London: RoutledgeFalmer

Falchikov, N. (2005) *Improving Assessment through Student Involvement*, London: RoutledgeFalmer

Falchikov, N. & Boud, D. (1989) Student self-assessment in higher education – a meta analysis, *Review of Educational Research* 59, pp. 395 - 430

Falchikov, N. & Goldfinch, J. (2000) Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks, Review of Educational Research, 70, pp. 287-322

Fay, B. (1996) *Contemporary Philosophy of Social Science*, Oxford: Blackwell

Forman, E. (1989) The role of peer interaction in the social construction of mathematical knowledge, *International Journal of Educational Research*, 13, pp. 55 – 69

Freeman, M. (1994) Peer assessment by groups of group work, *Assessment and Evaluation in Higher Education*, 20 (3) pp. 289 – 301

Fry, S. A. (1990) Implementation and evaluation of peer marking in higher education, *Assessment and Evaluation in Higher Education*, 15 (3) pp. 177-189

Gagne, M. & Zuckerman, M. (1999) Performance and learning goal orientation as moderators of social loafing and social facilitation, *Small Group Research,* 30, pp. 524 - 551

Garvin, J.W. & Butcher, A.C. (1995) Group projects for first-year university students: An evaluation, *Assessment and Evaluation in Higher Education*, 20 (3) pp. 273 – 289

Gatfield, T. (1999) Examining student satisfaction with group projects and peer assessment, *Assessment and Evaluation in Higher Education*, 24 (4) pp. 365 – 378

Gergen, K.J. (1995) Social Construction and the Educational Process, in Steffe, L. & Gale, J. (eds.) (1995) *Constructivism in Education*, Hillsdale, NJ: Lawrence Erlbaum

Gibbs, G. (1992) Down with Essays, *The New Academic*, Spring, p.18-19

Gibbs, G. (1992) Booklet 4: *Assessing More Students* – The teaching More Students Project 26 – 30, Oxford: The Polytechnics and Colleges Funding Council

Gibbs, G. (1992) Improving the quality of student learning through course design in R. Barnett (ed) *Learning to effect*, Milton Keynes: SRHE & Open University Press

Gibbs, G. (1995) *Assessing student-centred courses*, Oxford: OCSLD

Gibbs, G. (2006) How assessment frames student learning, in C. Bryan, & K. Clegg, (eds.) *Innovative Assessment in Higher Education*, London: Routledge

Gibbs, G. & Coffey, M. (2004) the impact of training university teachers on their teaching skills, their approach to teaching and the approach to learning of their students, *Active Learning in Higher Education,* 5, pp. 87 – 100

Gibbs, G. & Simpson, C. (2005) Conditions under which assessment supports student learning, *Learning and Teaching in Higher Education*, 1 (1) pp. 3 – 31

Gillies, R.M. (2000) The maintenance of cooperative and helping behaviours in cooperative groups, British Journal of Educational Psychology, 70, pp. 97 – 111

Gillies, R. & Ashman, A. (1996) Teaching collaborative skills to primary school children in classroom-based workgroups, *Learning and Instruction*, 6, pp. 187 – 200

Gillies, R.M. & Ashman, A. F. (2003) *Co-operative Learning*, London: RoutledgeFalmer

Glaser, B.G. & Strauss, A.S. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research, N*ew York: Aldine de Gruyter

Glasersfeld, E, von (1989) Learning as a Constructive Activity, in Murphy, P. & Moon, B (eds.) (1989) *Developments in Learning and Assessment*, Hodder & Stoughton

Goldfinch, J. & Raeside, R. (1990) Development of a peer assessment technique for obtaining individual marks on a group project, *Assessment and Evaluation in Higher Education*, 15, pp. 210-231

Goldfinch, J. (1994) Further development in peer assessment of group projects, *Assessment and Evaluation in Higher Education*, 19 (1) pp. 29 – 35

Grant, A. (1994) Group-project work: two enterprise case studies, in Thorley, L. & Gregory, R. (1994) *Group-based Learning in Higher Education*, London: Kogan Page, pp.127 – 135

Green, S. (1990) Analysis of Transferable Personal Skills Requested by Employers in Graduate Recruitment Advertisements in June 1989, Personal Skills Unit, University of Sheffield

Guba, E.G. and Lincoln, Y.S. (1988) Do Inquiry Paradigms Imply Inquiry Methodologies? in Fetterman, D. (1988) *Qualitative Approaches to Evaluation in Education,* Praegar

Guba, E.G. and Lincoln, Y.S. (1989) *Fourth Generation Evaluation*, Newbury Park CA: Sage

Guba, E.G. & Lincoln, Y.S. (1994) Competing paradigms in qualitative research, in N.K. Denzin & Y.S. Lincoln (eds.) *Handbook of Qualitative Research,* Thousand Oaks, CA: Sage

Guba, E.G. & Lincoln, Y.S. (2005) Competing paradigms in qualitative research, in N.K. Denzin & Y.S. Lincoln (eds.) *Handbook of Qualitative Research,* 3rd edn., Thousand Oaks, CA: Sage

Habeshaw, S., Gibbs, G. & Habeshaw, T. (1993) *53 Interesting Ways to Assess your students*, 3rd edn, Melksham: The Cromwell Press

Hall, K. (1995) Co-assessment: participation of students with staff in the assessment process. A report of work in Progress, paper given at the 2nd *European Electronic Conference on Assessment and Evaluation*, European Academic & Research Network (EARN); available at http://listserv.surfnet.nl/archives/earli-ae.html

Hanrahan, S.J. & Isaacs, G. (2001) Assessing Self- and Peer-Assessment: the students' views, *Higher Education Research and Development*, 20, (1)

Hardy, C.J. & Crace, R.K. (1991) The Effects of Task Structure and Teammate Competence on Social Loafing, *Journal of Sport and Exercise Psychology*, 13, pp. 372-381

Harkins, S.G. (1987) Social loafing and social facilitation, *Journal of Experimental Social Psychology*, 23, pp.1 – 18

Harkins, S.G. & Jackson, J.M. (1985) The role of evaluation in eliminating social loafing, *Personality and Social Psychology Bulletin*, 11 (4) pp. 457 - 465

Harkins, S.G. & Szymanski, K. (1989) Social loafing and group evaluation, *Journal of Personality and Social Psychology*, 56 (6) pp. 934 – 941

Harkins, S.G. & Petty, R.E. (1982) Effects of difficulty and task uniqueness on social loafing, *Journal of Personality and Social Psychology,* 43 (6) pp.1214 – 1229

Hattie, J. & Timperley, H. (2007) The power of feedback, *Review of Educational Research*, 77 (1) pp. 81 - 112

Healey, M. (1999) Using peer and self- assessment for assessing the contribution of individuals to a group project. Available at http://www.glos.ac.uk/gdn/abstracts/a69.htm Accessed 29 May 2007

Helms, M. & Haynes, P.J. (1990) When bad groups are good: An appraisal of learning from group projects, *Journal of Education for Business*, 66, pp.5 – 8

Henwood, K.L. & Pigeon, N.F. (1992) Qualitative research and psychological theorising, *British Journal of Psychology,* 83

Heron, J. (1981) Experiential Research Methodology, in Reason, P. and Rowan, J. (eds) *Human Inquiry*, New York, Wiley

Heron, J. (1981) Assessment revisited, in D.J. Boud (ed) *Developing Student Autonomy in Learning,* London: Kogan Page

Hertz-Lazarowitz, R. & Miller, N. (eds.) (1992) *Interaction in Cooperative Groups,* Cambridge: Cambridge University Press

Heywood, J. (2000) *Assessment in Higher Education*, London: Jessica Kingsley

Hinton, P. R., Brownlow, C., McMurray, I. & Cozens, B. (2004) *SPSS Explained*, London: Routledge

Hughes, I.E. & Large, B.J. (1993) Staff and peer group assessment of oral communication skills, *Studies in Higher Education*, 18 (3) pp.379 – 385

Jackson, J.M. & Harkins, S.G. (1985) Equity in Effort: An explanation of the social loafing effect, *Journal of Personality and Social Psychology*, 49, pp.1199 - 1206

Jackson, J.M. & Williams, K.D. (1985) Social loafing on difficult tasks: Working collectively can improve performance, *Journal of Personality and Social Psychology*, 49, pp. 937 - 942

Janesisk, V. (1994) The dance of qualitative research design, in Denzin, N. & Lincoln, Y. (eds) *Handbook of Qualitative Research,* London: Sage

Jaques, D. (2000) *Learning in Groups*, 3rd edn., London: Kogan Page

Johnson, K.P. (1993) *Team Peer Evaluations: A student-generated quantitative measurement of group membership performance.* Paper presented at the Annual Meeting of the Association for Education in Journalism and Mass Communication

Johnson, B., & Turner, L.A. (2003) in Tashakkori, A. & Teddlie, C. (2003) (2003) *Handbook of Mixed Methods*, London: Sage

Johnson, D.W. & Johnson, R.T. (1989) *Cooperation and Competition. Theory and Research*, Edina, MN: Interaction Book Company

Johnson, D.W. & Johnson, R.T. (1991) *Learning Together and Alone*, 4th edn., London: Allyn & Bacon

Johnson, D.W. & Johnson, R.T. (2003) Student motivation in co-operative groups, in Gillies, R.M. & Ashman, A. F. (2003) *Co-operative Learning*, London: RoutledgeFalmer

Johnson, D.W. & Johnson, F.P. (2003) *Joining Together, Group Theory and Group Skills*, 8th edn., London: Pearson

Johnson, D.W., Johnson, R.T. & Holubec, E.H. (1994) *The Nuts and Bolts of Cooperative Learning,* Edina, MN: Interaction Books

Johnson, D.W., Johnson, R.T. & Smith, K.A. (1991) *Active Learning: cooperation in the college classroom,* Edina, MN: Interaction Books

Johnson, D.W., Maruyama, G., Johnson, R., Nelson, D. & Skon, L. (1981) Effects of cooperative, competitive and individualistic goal structures on achievement: a meta-analysis, *Psychological Bulletin*, 89, pp. 47-62

Johnson, R. B. & Onwuegbuzie, A.J. (2004) Mixed methods research: a research paradigm whose time has come, *Educational Researcher*, 33, (7), pp.14 - 26

Johnson, S. (1999) The horrors of scientific research, *The Psychologist,* 12 (4)

Jordan, D. & LeMetais, J. (1997) Social skilling through cooperative learning, *Educational Research*, 39, pp3- 21

Karau, S. & Williams, K. (1993) Social loafing: a meta-analytic review and theoretical integration, *Journal of Personality and Social Psychology*, 65, pp. 681-706

Karau, S. & Williams, K. (1995) Social loafing: Research findings, implications and future directions, *Current Directions in Psychological Science*, 4 (5) pp.134 140

Kemmis, S. (1982) *The Action-Research Planner*, 2nd edn., Victoria: Deakin University Press

Kemmis, S. & McTaggart, R. (1988a) *The Action Research Reader*, 3rd edn., Victoria: Deakin University Press

Kemmis, S. & McTaggart, R. (1988b) *The Action Research Planner*, 3rd edn., Victoria: Deakin University Press

Kemmis. S. & Wilkinson, M. (1998) Participatory action research and the study of practice in B. Atweh, S. Kemmis & P. Weeks (eds.) *Action Research in Practice: Partnerships for social justice in education,* pp.21 – 36, New York: Routledge

Kerr, N.L. (1989) Illusions of efficacy: the effects of group size on perceived efficacy in social dilemmas, *Journal of Experimental Social Psychology*, 25, pp.287 - 313

Kerr, N.L. & Bruun, S. (1983) Dispensability of Member Effort and Group Motivation Losses: Free-Rider Effects, *Journal of Personality and Social Psychology*, 44 (1) pp.78 – 94

Kerr, N.L. & Stanfel, J.A. (1993) Role schemata and member motivation in task groups, *Personality and Social Psychology Bulletin*, 19, pp. 432 - 444

Kilic, G.B. & Cakan, M. (2006) The analysis of the impact of individual weighting factor on individual scores, *Assessment and Evaluation in Higher Education*, 31 (6) pp. 639 – 654

Klenowski, V. (1995) Student self-evaluation processes in student-centred teaching and learning contexts of Australia and England, *Assessment in Education*, 2 (2) pp. 145 163

Knight, J. (2004) Comparison of student perception and performance in individual and group assessments in practical classes, *Journal of Geography in Higher Education,* 28 (1) pp. 63 – 81

Kniveton, B.H. (1996) Student perceptions of assessment methods, *Assessment and Evaluation in Higher Education*, 21, 3

Kohlberg, L. (1969) Stage and sequence: the cognitive developmental approach to socialization, in D.A. Goslin (ed.) *Handbook of Socialisation Theory and Research*, New York: Rand McNally

Kolb, D. (1984) *Experiential Learning*, New Jersey: Prentice-Hall

Krause, J.E. & Popovich, N.G. (1996) A group interaction peer/self assessment process in a pharmacy practice course, *American Journal of Pharmaceutical Education*, 60, pp136 – 145

Kugihara, N. (1999) Gender and Social Loafing in Japan, *Journal of Social Psychology*, 139 (4) pp. 516-526.

Kwan, K. & Leung, R. (1996) Tutor versus peer group assessment of student performance in a simulation training exercise, *Assessment and Evaluation in Higher Education,* 21, pp. 205 214

LaFasto, F. & Larson C. (2001). *When Teams Work Best.* Thousand Oaks, CA: Sage

Langan, A.M., Wheater, C.P. & Shaw, E.M. (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) pp.21-34

Latané, B., Williams, K. & Harkins, S. (1979) 'Many hands make light the work: the causes and consequences of social loafing', *Journal of Personality and Social Psychology*, 37, pp. 822 -832

Lazarowitz, R. (1991) Learning biology cooperatively: An Israeli junior high school study, *Cooperative Learning*, 11 (3) pp.19 - 21

Lazarowitz, R. & Karsenty, G. (1990) Cooperative learning and students' achievement, process skills, learning environment, and self esteem in tenth grade biology classrooms, in S. Sharan (ed.) *Cooperative Learning Theory and Research,* pp.123 – 149, New York: Praeger

Lejk, M. & Wyvill, M. & Farrow, S. (1996) A survey of methods of deriving individual grades from group assessments, *Assessment and Evaluation in Higher Education,* 21 (3) pp. 267 - 280

Lejk, M. & 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) pp.61 – 72

Lejk, M. & 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) pp. 551 - 561

Lejk, M. & 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, (6)

Levin, P. (2005) *Successful Teamwork*, Maidenhead: Open University Press

Lewin, K. (1948) *Resolving Social Conflicts*, New York: Harper

Lewin, K. (1951) *Field Theory in Social Science*, New York: Harper

Lewin, K. (1952) *Field Theory in Social Science*, London: Tavistock Publications

Lewin, K. (1946) Action research and minority problems, in Kemmis, S. & McTaggart, R. (1988a) *The Action Research Reader*, 3rd edn., Victoria: Deakin University Press

Li, L.K.Y. (2001) Some refinements on peer assessment of group projects, *Assessment and Evaluation in Higher Education*, 26, (1)

Lincoln, Y.S. & Guba, E.G. (1985) *Naturalistic Inquiry*, Beverly Hill: CA: Sage

Lincoln, Y.S. & Guba, E.G. (2000) Paradigmatic controversies, contradictions and emerging confluences, in Denzin, N. K. & Lincoln, Y.S. (2005) *Handbook of Qualitative Research,* 3rd edn., London: Sage

Livingstone, D. & Lynch, K. (2000) Group project work and student-centred active learning: two different experiences, *Studies in Higher Education*, 25 (3) pp. 325 - 345

Loevinger, J. (1966) The meaning and measurement of ego-development, *American Psychologist*, 21, pp.195 – 206

Magin, D.J. (2001) A novel way for comparing the reliability of multiple peer assessments with that of single teacher's assessment of group process work, *Assessment and Evaluation in Higher Education*, 26 (2)
Magin, D.J. (2001) Reciprocity as a source of bias in multiple peer assessment of group-work, *Studies in Higher Education*, 26, pp. 53-63
Maguire, S. & Edmondson, S. (2001) Student evaluation and assessment of group projects, *Journal of Geography in Higher Education*, 25 (2) pp. 209 - 217
Marshall, C. & Rossman, G. (1989) *Designing Qualitative Research*, London: Sage
Marshall, C. & Rossman, G. (1999) *Designing Qualitative Research*, 3rd edn., London: Sage
Martens, R. & Dochy, F. (1997) Assessment and feedback as student support devices, *Studies in Educational Evaluation*, 23, pp. 257 – 275
Mason, J. (1996) *Qualitative Researching*, London: Sage
McAuley, T. (2009a) *Students' perceptions of peer feedback on contribution to assessed group work,* Paper presented at the Oxford Brookes University Teaching and Learning Conference
McAuley, T. (2009b) *Assessment and extrinsic learners: the tail that wags the dog?* Paper presented at the Westminster Institute of Education Teaching and Learning Conference
McDowell, L. (1995) The impact of innovative assessment on student learning, *Innovations in Education and Training International*, 32, pp.302 – 313
McDowell, L. & Harman, K. (2008) 'The designer', 'the teacher' and 'the assessor': changing academic identities. Paper presented at the Society for Research into Higher Education Conference, December, 2008
McKernan, J. (1996) *Curriculum Action Research*, 2nd edn., London: Kogan Page
McNiff, J. (1988) *Action Research: Principles and Practice,* London: Routledge
McNiff, J. (2002) *Action Research: Principles and Practice*, London: RoutledgeFalmer

McNiff, J. & Whitehead, J. (2006) *All You Need To Know About Action Research*, London: Sage

McWhaw, K., Schnackenberg, H., Sclater, J. & Abrami, P.C. (2003) Helping students become collaborative learners in R.M. Gillies & A.F. Ashman, *Co-operative Learning*, London: RoutledgeFalmer

Mehan, H. (1979) *Learning lessons: Social Organisation in the Classroom*, Cambridge, MA: Harvard University Press

Mehmetoglu, M. & Altinay, L. (2006) Examination of grounded theory analysis with an application to hospitality research, *Hospitality Management*, 12, pp. 12 - 33

Mello, J.A. (1993) Improving individual member accountability in small work group settings, *Journal of Management Education*, 17 (2) pp.253 – 259

Mesch, D. (1991) The jigsaw technique: A way to establish individual accountability in group work, *Journal of Management Education*, 15 (3) pp. 335-358

Meyers, S.A. (1997) Increasing student participation and productivity in small-group activities for psychology classes, *Teaching of Psychology*, 24 (2) pp. 105 - 115

Miles, M.B. & Huberman, A.M. (1984) *Qualitative Data Analysis: A sourcebook of new methods*, Beverly Hills, Ca: Sage

Miles, M.B. & Huberman, A.M. (1994) *Qualitative Data Analysis; An Expanded Sourcebook,* 2nd edn., London: Sage

Millar, J. (2008) *Understanding Feedback*, ASKE Strand 2 Project Report, Oxford: Oxford Brookes University

Miller, P.J. (2003) The effect of scoring criteria specificity on peer and self assessment, *Assessment and Evaluation in Higher Education*, 28 (4) pp. 383 – 394

Morgan, B.M. (2003) Cooperative learning in higher education: undergraduate student reflections on group examinations for group grades – a study of the reflections of 140 university seniors who participated in cooperative written examinations for group grades , *College Student Journal*, retrieved on 26 September 2006 from: http://www.findarticles.com/p/articles/mi_m0FCR/is_1_37/ai_99816478

Morgan, D.L. (2007) Paradigms lost and pragmatism regained: methodological implications of combining qualitative and quantitative methods, *Journal of Mixed Methods Research*, 1, pp. 48 – 76

Mowl, G. & Pain, R. (1995) Using self and peer assessment to improve students' essay writing: a case from geography, *Innovations in Education and Teaching International,* 32 (4) pp324 – 335

Neuman, W.L. (2004) *Basics of Social Research*, London: Pearson

Nicol, D. & Macfarlane-Dick, D. (2007) *Rethinking formative assessment in HE: a theoretical model and seven principles of good feedback practice* York: HEA. [Online] Retrieved 7 September, 2007 from: www.ltsn.ac.uk/genericcentre

North, A.C., Linley, P.A. & Hargreaves, D.J. (2000) Social loafing in a cooperative classroom, *Educational Psychology*, 20 (4) pp. 389 - 392

O'Donovan, B., Price, M. & Rust, C. (2001) The student experience of criterion-referenced assessment through the use of a common criteria assessment grid, *Innovations in Learning and Teaching International*, 38, (1) pp. 74 – 85

O'Donovan, B., Price, M. & Rust, C. (2008) Developing student understanding of assessment standards: a nested hierarchy of approaches, *Teaching in Higher Education* 13 (2) pp. 205-218

O'Hanlon, C. (ed.) (1996) *Professional Development through Action Research in Educational Settings,* London: The Falmer Press

Oldfield, K.A. & Macalpine, J & Mark, K. (1995) Peer and Self-Assessment at Tertiary Level an Experiential Report, *Assessment and Evaluation in Higher Education,* 20 (1)

Orpen, C. (1982) Student versus lecturer assessment of learning: a research note, *Higher Education*, 11, pp. 567 - 572

Orsmond, P. & Merry, S. (1996) The importance of marking criteria in the use of peer assessment, *Assessment and Evaluation in Higher Education,* 21 (3)

Orsmond, P., Merry, S. & Reiling, K. (1997) A study in self-assessment: tutor and students' perceptions of performance criteria, *Assessment and Evaluation in Higher Education,* 22 (4)

Orsmond, P., Merry, S. & Reiling, K (2000) The use of student derived marking criteria in peer and self-assessment, *Assessment and Evaluation in Higher Education*, 25 (1) pp. 23 – 37

Orsmond, P., Merry, S. & Reiling, K. (2002) The use of exemplars and formative feedback when using student derived marking criteria in peer and self-assessment, *Assessment and Evaluation in Higher Education*, v 27 (4)

Otter, S. (1995) Assessing competence: the experience of the Enterprise in Higher Education Initiative in, A. Edwards & P. Knight (eds.) *Assessing Competence in Higher Education,* London: Kogan Page

Pain, H., Bull, S. & Brna (1996) A student model "for its own sake" available at http://cbl.leeds.ac.uk/~paul/papers/euroaiedpapers96/smpaper/smpaper.html

Patton, M. (2002) *Qualitative Research and Evaluation Methods*, 3rd edn., London: Sage

Perry, W.G. (1999) *Forms of Intellectual and Ethical Development in the College Years,* San Francisco: Jossey-Bass

Peters, M. (1996) Student attitudes to alternative forms of assessment and to openness, *Open Learning*, 11, pp. 48 – 50

Phipps, M., Phipps, C., Kask, S. & Higgins, S. (2001) University students' perceptions of cooperative learning: implications for administrators and instructors, *The Journal of Experiential Education*, 24 (1) pp.14 – 21

Piaget, J. (1965) *The Moral Judgement of the Child*, London: Routledge & Kegan Paul

Polanyi, M. (1962) *Personal Knowledge: towards a post-critical philosophy*, London: Routledge and Kegan Paul

Polanyi, M. (1998) The tacit dimension, in: L. Prusak, Ed) *Knowledge in Organisation*: Boston: Butterworth Heinemann

Pollard, A. (2002) *Reflective Teaching*, London: Continuum

Pond, K., ul-Haq, R. and Wade, W. (1995) Peer review: a precursor to peer assessment, *Innovations in Education and Training International*, 32, pp. 314 – 323

Pond, K. & ul-Haq, R. (1997) Learning to assess students using peer review, *Studies in Educational Evaluation*, 23 (4) pp. 331 - 348

Price, K.H. & Harrison, D.A. (2006) Witholding inputs in team context: member composition, interaction process, evaluation structure, and social loafing. *Journal of Applied Psycology*, 91 (6) pp.1375 - 1384

Price, M. & O'Donovan, B. (2006) Understanding criteria and feedback, in C. Bryan, & K. Clegg, (eds.) *Innovative Assessment in Higher Education*, London: Routledge

Price, M., O'Donovan, B., Rust, C., Carroll, J. & Gibbs, G. (2008) *ASKe position paper on Assessment,* Oxford Brookes University:CETL

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

Race, P. (1998) Practical points on peer-assessment, in Brown, S. (ed) Peer Assessment in Practice, Birmingham: SEDA

Rafiq, Y. & Fullerton, H. (1996) Peer assessment of group projects in civil engineering, *Assessment and Evaluation in Higher Education*, 21, (1) pp. 69 – 82

Ramsden, P. (1992) *Learning to Teach in Higher Education*, London: Routledge

Ravenscroft, S. P., Buckless, F.A. & Hassall, T. (1999) Cooperative Learning – a literature guide, *Accounting Education* 8 (2) pp. 163-176

Reilly, R.R. Smither, J.W. & Vasilopoulis, N.L. (1996) An longitudinal study of upward feedback, *Personnel Psychology*, 48, pp.599 - 612

Reimann, P. & Spada, H. (1996) *Learning in Humans and Machines*, Oxford: Elsevier

Ritter, L. (1998) Peer assessment: lessons and pitfalls in S. Brown (ed.) *Peer Assessment in Practice*, Birmingham: SEDA

Robson, C. (1993) *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*, Blackwell

Ross, J, & Rolheiser, C. (2003) Assessment in co-operative learning, in, R.M. Gillies & A. F. Ashman (2003) *Co-operative Learning*, London: RoutledgeFalmer

Rust, C. (2000) An opinion piece – a possible student-centred assessment solution to some of the current problems of modular degree programmes, *Active Learning in Higher Education*, 1, pp.126 - 131
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)
Rust, C., O'Donovan, B. & Price, M. (2005) A social constructivist assessment process model: how the research literature shows us this could be best practice, *Assessment and Evaluation in Higher Education*, 30 (3) pp. 231 – 240
Ryan, G. (1997) Ensuring that students develop an adequate and well-structured knowledge base, in D. Boud and G. Feletti (eds.) *The Challenge of Problem Based Learning,* London: Kogan Page
Sadler, R. (1989) Formative assessment and the design of instructional systems, *Instructional Science*, 18, pp. 119-144
Sadler, D.R. (1998) Formative assessment: revisiting the territory, *Assessment in Education*, 5 (1) pp.77 - 84
Sambell, K. & McDowell, L. (1997) The value of self and peer assessment to the developing lifelong learner in C,. Rust (ed.) *Improving Student Learning – improving students as learners*, pp. 56 – 66, Oxford: OCSLD
Sambell, K., McDowell, L. & Brown, S. (1997) "But is it fair?" : an exploratory study of student perceptions of the consequential validity of assessment, *Studies in Educational Evaluation*, 23, pp. 349 – 371
Santrock, J.W. (2001) *Educational Psychology*, New York: McGraw Hill
Schachar, H. (2003) Who gains what from co-operative learning: an overview of eight studies, in Gillies, R.M. & Ashman, A. F. (2003) *Co-operative Learning*, London: RoutledgeFalmer
Schachar, H. & Sharan, S. (1994) Talking, relating and achieving: Effects of cooperative learning and whole class-instruction, *Cognition and Instruction*, 12, pp. 313 – 353
Schechtman, Z. (1991) A revised group assessment procedure for predicting initial teaching success, *Educational and Psychological Measurement*, 51 (4) pp. 963 - 974

Schechtman, Z. (1992) Interrater reliability of a single group assessment procedure administered in several educational settings, *Journal of Personnel Evaluation in Education,* 6, pp. 31 - 39

Schechtman, Z. & Godfried, L. (1993) Assessing the performance and personal traits of teacher education student by a group assessment procedure: a study of concurrent and construct validity, *Journal of Teacher Education*, 44 (2) pp. 130 – 138

Schmeck, R.R. (ed.) (1988) *Learning Strategies and Learning Styles*, London: Plenum Press

Schnake, M.E. (1991) Equity in Effort: The 'sucker effect' in co-acting groups. *Journal of Management*, 17(1), pp. 41-55.

Schön, D.A. (1991)*The Reflective Practitioner: How Professionals Think in Action,* Aldershot: Avebury

Schwandt, T.A. (2000) Three epistemological stances for qualitative inquiry: Interpretivism, hermeneutics, and social constructionism, in Denzin, N. K. & Lincoln, Y.S. *Handbook of Qualitative Research,* London: Sage

Scoffield, S & Brindley, C. (1998) Peer assessment in undergraduate programmes, *Teaching in Higher Education*, 3 (1) pp. 79 – 90

Seale, C. (1999) *The Quality of Qualitative Research*, London: Sage

Searby, M., Ewers, T., (1997) An evaluation of the use of peer assessment in higher education: a case study in the school of music, *Assessment and Evaluation in Higher Education*, 24 (4)

Seidman, I. (1998) *Interviewing as Qualitative Research*, 2nd edn., London: Teachers' College Press

Sharan, S. (1980) Cooperative learning in small groups: Recent methods and effects on achievement, attitudes, and ethnic relations, *Review of Educational Research,* 50, pp. 241 – 272

Sharan, S. (2003) Large classes, small groups: a social systems approach, in Gillies, R.M. & Ashman, A. F. (2003) *Co-operative Learning*, London: RoutledgeFalmer

Sharan, S. & Schachar, H. (1988) *Language and Learning in the Cooperative Classroom*, New York: Springer-Verlag

Sharan, S. & Shaulov, A. (1990) Cooperative learning, motivation to learn, and academic achievement in S. Sharan (ed.) *Cooperative Learning Theory and Research* (pp.123 – 149) New York: Praeger

Sharan, S. & Sharan, Y. (1976) *Small-group Teaching*, Englewood Cliffs, NJ: Educational Technology

Sharp, S. (2006) Deriving individual student marks from a tutor's assessment of group work, *Assessment and Evaluation in Higher Education*, 31 (3) pp. 329-343

Sheppard, J.A. & Taylor, K.M. (1999) Social loafing and expectancy-value theory, *Personality and Social Psychology Bulletin*, 25, pp.1147 - 58

Sherif, M. & Sherif, C. (1956) *An Outline of Social Psychology*, New York: Harper & Row

Silverman, D. (2001) *Interpreting Qualitative Data*, 2nd edn., London: Sage

Sivan, A. (2000) The implementation of peer assessment: an action research approach, *Assessment in Education*, 7 (2) pp. 193 – 209

Slavin, R.E. (1989a) Cooperative learning and student achievement: six theoretical perspectives, *Advances in Motivation and Achievement: Motivation Enhancing Environments*, 6, pp. 161 – 177

Slavin, R.E. (1989b) Research on cooperative learning: consensus and controversy, *Educational Leadership*, 47, (4) pp. 52 – 55

Slavin, R.E. (1992) When and why does cooperative learning increase achievement? In Hertz-Lazarowitz, R. & Miller, N. (eds.) (1992) *Interaction in Cooperative Groups,* Cambridge: Cambridge University Press

Slavin, R., Sharan, S., Kagan, S., Hertz-Lazarowitz, R., Webb, C. & Schmuck, R. (1985) *Learning to Cooperate and Cooperating to Learn*, London: Plenum Press

Smith, J.K. (1984) The problem of criteria for judging interpretive inquiry, *Educational Evaluation and Policy Analysis*, 6, pp. 379 - 391

Smith, M. (1945) Social situation, social behavior and social group, *Psychological Review*, 52, pp. 224-229

Smith, M. K. (1996; 2001) 'Action research', *the encyclopaedia of informal education*, www.infed.org/research/b-actres.htm retrieved on 10 November 2006

Smither, J.W., London, M., Vasilopoulis, N.L., Reilly, R.R., Millsap, R.E. & Salvemini, N. (1995) An examination of the effects of an upward feedback program over time, *Personnel Psychology*, 48 (1) pp.1 - 32

Stanier, L. (1997) Peer assessment and group work as vehicles for student empowerment: a module evaluation, *Journal of Geography in Higher Education*, 21 (1) pp. 95 - 99

Stefani, L.A.J. (1992) Comparison of collaborative, self, peer and tutor assessment in a biochemistry practical, *Biochemical Education*, 20, pp. 148 - 151

Stefani, L.A.J. (1994) Peer, self and tutor assessment: relative reliabilities, *Studies in Higher Education*, 19 (1) pp.69 – 75

Stefani, L.A.J. (1998) Assessment in partnership with learners, *Assessment and Evaluation in Higher Education*, 23 (4) pp. 339 – 350

Steffe, L. & Gale, J. (eds.) (1995) *Constructivism in Education*, Hillsdale, NJ: Lawrence Erlbaum

Strachan, I.B. & Wilcox, S. (1996) Peer and self assessment of group work: developing an effective response to increased enrolment in a third year course in microclimatology, *Journal of Geography in Higher Education*, 20, pp. 343 – 353

Strauss, A. & Corbin, J. (1998) *Basics of Qualitative Research*, 2nd edn., London: Sage

Struyven, K., Dochy, F. & Janssens, S. (2005) Students' perceptions about evaluation and assessment in higher education: a review, *Assessment and Evaluation in Higher Education* 30 (4) pp. 325 – 341

Swann, J. & Pratt, J. (1999) *Improving Education: Realist Approaches to Method and Research*, London: Cassell

Tashakkori, A. & Teddlie, C. (1998) *Mixed Methodology: Combining Qualitative and Quantitative Approaches*, London: Sage

Tashakkori, A. & Teddlie, C. (eds.) (2003) *Handbook of Mixed Methods*, London: Sage

Thorley, L. & Gregory, R. (1994) *Group-based Learning in Higher Education*, London: Kogan Page

TMP Worldwide Research, (1998) *Soft Skills, Hard Facts,* London: TMP Worldwide Research

Tolmie, A.K. & Warden, D.A. (eds.) (1994) *Group and Interactive Learning*, Southampton: Computational Mechanics Publications

Topping, K. (1998) Peer assessment between students and in colleges and universities, *Review of Educational Research*, 68, pp. 249 - 276

Topping, K.J., Smith, E.F., Swanson, I. & Elliot, A. (2000) Formative peer assessment of academic writing between postgraduate students, *Assessment and Evaluation in Higher Education*, 25 (2)

Torbert, B. (2004) *Action Inquiry*, San Francisco: Berrett-Koehler

Tyler, T. (2008) Personal correspondence

Underwood, J.D.M. (2003) Student attitudes towards socially acceptable and unacceptable group working practices, *British Journal of Psychology*, 94, pp. 319 – 337

Vygotsky, L.S. (1962) *Thought and Language*, Cambridge: MIT Press

Vygotsky, L.S. (1978) *Mind in Society: the development of higher psychological processes,* London: Harvard University Press

Webb, N. (1985) Student interaction and learning in small groups: a research summary, in R. Slavin, S. Sharon, S. Kagan, R. Hertz-Larowitz, C. Webb & R. Schmuck (eds.) *Learning to cooperate, cooperating to learn* (pp.5 – 15) New York: Plenum

Webb, N. (1989) Peer interaction and learning in small groups, *International Journal of Educational Research,* 13, pp. 21 – 39

Webb, N. (1991) Task related verbal interaction and mathematical learning in small groups, *Journal for Research in Mathematics Education*, 22, pp. 366-389

Webb, N. (1992) Testing a theoretical model of student interaction and learning in small groups in R. Hertz-Lazarowitz, & N. Miller (eds.) (1992) *Interaction in Cooperative Groups,* Cambridge: Cambridge University Press

Webb, N. & Farivar, S. (1994) Promoting helping behaviour in cooperative small groups in middle school mathematics, *American Educational Research Journal*, 31, pp. 369 – 395

Webb, N., Troper, J., & Fall, R. (1995) Constructive activity and learning in collaborative small groups, *Journal of Educational Psychology*, 87, pp. 406 – 423

Webster, F., Pepper, D. & Jenkins, A. (2000) Assessing the undergraduate dissertation, *Assessment and Evaluation in Higher Education*, 25, (1) pp.71 - 80

Webster, F., Pepper, D. & Jenkins, A. (2000) Assessing the undergraduate dissertation, *Assessment and Evaluation in Higher Education*, 25 (1) pp.72 – 80

Whitehead, J. & McNiff, J. (2006) *Action Research Living Theory*, London: Sage

Williams, E. (1992) Students attitudes towards approaches to learning and assessment, *Assessment and Evaluation in Higher Education*, 17 pp.45 – 58

Williams, K.D. & Karau, S.J. (1991) Social loafing and social compensation: The effects of expectations of co-worker performance, *Journal of Personality and Social Psychology,* 61, pp. 570 – 581

Williams, K. D., Harkins, S.G. & Latane, B. (1981) Identifiability as a deterrent to social loafing, *Journal of Personality and Social Psychology*, 40, pp. 303 – 311.

Willig, C (2001) *Qualitative Research in Psychology*, Maidenhead: Open University Press

Winne, P.H. & Butler, D.L. (1994) Student cognition in learning from teaching, in T. Husen & T. Postlewaite (eds.) *International Encyclopaedia of Education*, 2nd edn., Oxford Pergamon

Winter, R. (1989) *Learning from Experience*, Lewes: The Falmer Press

Wittrock, M. (1990) Generative processes of comprehension, *Educational Psychologist*, 24, pp. 345 - 376

Yorke, M (2003) Formative assessment in higher education: moves towards theory and the enhancement of pedagogic practice, *Higher Education*, 45, pp. 477 - 501

Zajonc, R.B. (1965) Social facilitation, *Science*, 149, pp. 269 - 274

Zuber-Skerritt, O. (1991) *Professional Development in Higher Education: A theoretical framework for Action Research*, Brisbane: Centre for the Advancement of Learning and Teaching

Zuber-Skerritt, O. (1992) Action *Research in Higher Education*, London: Kogan Page

Appendix 1: Elements of reflecting, planning and acting in each of the three action research cycles

Reflection leading to research questions	Planning, acting and reflecting elements of action research process incorporating literature reviews and data collection/analysis	Result(s) of action
Cycle 1: Initial research question		
1. Reflecting: What are the advantages and disadvantages of group work and what are the theories that underpin it? Chapter 2	Planning: Literature Review 1: dis/advantages of group work and theories underpinning the use of group work	
2. Reflecting: What are students' perceptions of group work? What do they like about it? What problems do they report? Chapter 4	Questions on module evaluation forms to find out students' perceptions of group work Acting: Data collection and analysis (Questionnaire 1) Reflecting: on findings and comparing them with the literature on students' perceptions of group work (Literature Review 2)	Result: Main problem identified as the practice of awarding group members the same mark irrespective of individual contribution (confirmed by literature review). Students want a way to identify individual contribution to assessed group work to deter or identify "social loafers".
3. Reflecting: Why does social loafing occur? Chapter 5: Section 1	Planning: Literature review 3 (what does the literature say about why social loafing occurs?) Acting: (carrying out literature review) Reflecting: on literature review	Result: Factors identified to help alleviate the problem of social loafing. Main factor identified: importance of "identifiability."
4. Reflecting: How can identifiability be incorporated into the assessment of group work? What strategies of identifying individual contribution to group work exist? Chapter 5: Section 2	Planning: Literature review 4 (what does the literature say about methods of identifying individual contribution to group work?) Acting: (carrying out literature review) Reflecting: on literature review	Result: Options identified in the literature.
Cycle 2: Choices presented to students to inform decisions		
5. Reflecting: Which method would (my) students choose and why? Chapter 6	Planning how to present intervention choices to students Acting: Data collection and analysis (Questionnaires 2 and 3)	Result: Students decide on choice of strategy: peer assessment of contribution to group effort.

6. Reflecting: What factors need to be taken into account when implementing peer assessment? What does the literature advise on criterion-referenced versus holistic assessment? If the former, who should choose the criteria? Should self-assessment be included or excluded? Should the allocation of summative marks be kept confidential or not? How could a formative element be incorporated into the assessment process so students benefit from peer feedback? Chapter 7	Planning, acting and reflecting on: Literature review 5 (what factors need to be taken into account when putting in place a peer assessment mechanism? What are the advantages of peer assessment?)	**Result:** Factors which need to be taken into account identified in the literature.
7. Reflecting on student choices: Would students choose criterion-referenced or holistic peer assessment? If the former what criteria would they choose? Chapter 8	Planning: questions to elicit students' responses to factors identified e.g. (a) criterion-referenced or holistic assessment? (b) % of marks to be allocated?, (c) choice of criteria Acting: Data collection and analysis (Questionnaires 4, 5 & 6) and reflecting on findings Planning the intervention based on reflecting on findings Acting: implementing the intervention	**Result:** Students' responses to the factors identified in the literature ascertained. **Result:** Summative and formative peer assessment of contribution to group product implemented.
Cycle 3: Students' evaluation of intervention		
8. Reflecting: What are students' perceptions of a formative and summative peer assessment mechanism introduced to identify individual students' contributions to a group assignment? Does such a mechanism alleviate social loafing? Chapter 9	Planning: how the students are to evaluate the strategy Acting: implementing the evaluation; data collection and analysis (Questionnaires 7 & 8) Reflecting: on students' evaluation of the process and my own learning	**Result:** Students' evaluation of formative and summative mechanism to assess peers' contribution to assessed group work.

Appendix 2: Summative Peer and Self-Assessment Form

One of the advantages of working as a member of a group is that you can all benefit from the strengths of each other. The purpose of this exercise is to give recognition to the varied contributions that individuals made to the group during this module. The marks you assign will remain confidential.

Indicate your agreement with the extent to which each member of your group met each criterion using the scale below:

Strongly agree	Agree	Partially agree		Partially disagree		Disagree	Strongly disagree
7	6	5	4	4	3	2	1

	You	Member 2	Member 3	Member 4
Insert names:				
Excellent quality of contribution (written work and ideas)				
Very respectful of others (listening to and respecting others' opinions and ideas)				
Excellent time management (meeting deadlines; being punctual)				
Very cooperative (supporting others; willingness to compromise; flexibility)				
Excellent commitment (attendance; preparation; effort; enthusiasm)				
Reliability (doing what they say they would do)				
Excellent organisation skills (organising others, meetings, content, resources)				
Total				

Overall assessment of contribution:

You have 100 marks to divide between group members. Based on the totals above, how many marks would you assign each individual, including yourself? (The total must add up to 100.)				

Any additional information you think the module leader should know:

Appendix 3: Peer Feedback Form

Name of student to whom you are providing feedback: Your name:

During this module you have worked as part of a small group to produce a poster and understand semiotics.

At the beginning of the module you decided on the criteria you wanted to use to evaluate each member of your group's contribution to the above activities:

You chose the criteria below:

- **Quality of contribution** (written work and ideas during discussion)
- **Respect for others** (listening to and respecting others' opinions and ideas)
- **Time management** (meeting deadlines; being punctual)
- **Cooperation** (supporting others; willingness to compromise; flexibility)
- **Commitment** (attendance; motivation; effort; enthusiasm)
- **Reliability** (doing what you say you would do)
- **Organisation skills** (organising others, meetings, content, resources)

Using a separate sheet for each member of your group, please give them feedback by writing some comments, based on the above criteria, on their contribution to your group during this module. You should include a form assessing yourself too. Make sure your comments are typed and:

- Describe behaviour clearly giving specific examples
- Describe what you think was good about the behaviour
- Describe what behaviour you think needs to be improved
- Give suggestions as to how behaviour could be improved

Quality of contribution:
Respect for others:

Time management:
Cooperation:
Commitment:
Reliability:
Organisation skills:
Additional comments:

Appendix 4: Questionnaires

Questionnaire 1

During this module you have had the opportunity to work in groups to produce a poster on the elements of communication and to discuss a semiotic analysis of a text after which you wrote individual essays.

What did you like best about working as part of your group during this module?

What did you like least about working as part of your group during this module?

Questionnaire 2

	1 = not			4 = very
You undertake a group project. The tutor awards one mark for the project and all group members get this same mark. How effective would this method be in ensuring IA?	1	2	3	4
Please give reasons your answer:				

	1 = not			4 = very
You undertake a group project. The group mark is then weighted according to summative peer assessment of contribution to group effort based on criteria agreed upon at the beginning of the module. How effective would this method be in ensuring IA?	1	2	3	4
Please give reasons your answer:				

Questionnaire 3

In this module you will be working in a small group to complete tasks in class as well as an assessed group assignment. There are various ways to assess group work; the box below gives three options.

4. **Peer assessment of contribution to group effort.** The tutor gives a mark for the group project, which is then modified according to peer assessment of contribution to the group effort. Peer assessment marks are awarded based on criteria agreed on by the group at the beginning of the module. The marks awarded by peers are averaged to arrive at a mark and remain confidential.

5. **Distribution of a pool of marks.** Students receive a group mark (for example, 60%) for the project from the tutor. The group mark is then multiplied by the number of members in the group (for example, three members in a group= 60 x 3 = 180 marks.) Students then discuss and agree how to divide the 180 marks between themselves, for example: student A might get 70%, student B 60% and student C 50%.

6. **Equally shared mark with exceptional tutor intervention.** Each member of the group is awarded the same group mark unless the tutor is approached because one or more students are felt not to be contributing. A meeting between the tutor and all members of the group takes place should this happen, the problems are discussed and individual marks agreed between the group and tutor.

Please indicate which of the options you would prefer by circling 1, 2 or 3 above.

In the space below, please give reasons for your choice:

Questionnaire 4

Criterion-referenced (CR) peer assessment involves summative assessment against agreed criteria. So, for example, you would assess each group member against a number of agreed criteria (such as contributions to discussions, attendance at meetings, etc.) negotiated at the beginning of the module.

Holistic (H) peer assessment would involve giving each group member a grade/ mark based on an intuitive, general 'gut' feeling about what each member deserves. It would involve one mark only and there would be no criteria on which to base your judgement.

Which of the above methods do you think would be more valid and why?

Questionnaire 5

Currently 10% of the module mark is based on peer assessment of contribution. Do you think this % is too low, too high or about right? Please circle your answer below:

1. Too low

2. Too high

3. About right

Please give reasons for your answer:

Questionnaire 6

In your groups you have had time to discuss what makes a really good group member. Now, as a group, list the five most important behaviours/skills of a good group member in the space below:

1.

2.

3.

4.

5.

Questionnaire 7

During this module you have been involved in setting criteria for assessing each other's contribution to group work.

Do you think it is a good idea to be involved in setting criteria for assessment? (Please circle one answer below.)

Yes No Not sure

In the space below, please give reasons for your answer:

Questionnaire 8

There were two elements in the peer assessment process. The first part included providing qualitative feedback (writing comments to help your peers consider their group work skills and contribution). The second part involved providing marks based on agreed criteria such as listening skills.

1. How did you feel about writing feedback to help other members of your group see what they were good at and what group skills they needed to improve?

2. Would you have preferred to give your feedback anonymously?

Yes No

3. Please give a reason for your answer.

4. On most modules when you work in groups, you do not get any feedback on your group working skills. Do you think it is a good idea to receive feedback from other group members on your contribution and group working skills?

Yes No Not sure

5. Please give reasons for your answer.

6. To what extent did you find the feedback from your group members about your group working skills beneficial?

Very beneficial Quite beneficial Not very beneficial Not at all beneficial

7. Please give reasons for your answer:

8. The reason you were asked to assess other group members' contributions is because the tutor is not present when you meet outside class and is, therefore, not in a position to see what sort of contributions you make. How did you feel about the second part of the process, i.e. giving each member of your group marks based on the criteria you had decided on at the beginning of the module?

9. The marks you awarded were anonymous. Do you think the fact that they were anonymous was a good idea?

Yes No Not sure

10. Please give reasons for your answer.

11. How well did your group work together in general?

Very well Well Satisfactorily Not too well Poorly

12. To what extent would you like to work in the same group on another module?

Very much A fair amount Don't mind Not really Not at all

13. What were the main things you learned from working in a group on this module?

14. Are you male or female?

Male Female

Any other comments: