University tutors' beliefs about and practices in assessing undergraduates' writing - A New Zealand case study

A thesis

submitted in full fulfilment

of the requirements for the degree of

Doctor of Philosophy

in

Applied Linguistics

at

The University of Waikato

by

JINRUI LI

_____

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2012

STATEMENT OF INTELLECTUAL OWNERSHIP

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgement has been made.

Signature:

Date: March 30th, 2012

ABSTRACT

Although teacher cognition has been explored widely, university tutor cognition of professional activities, such as evaluating and giving written feedback on students' written work, has rarely been explored. Very few studies on teacher cognition of giving feedback have included data of real practice collected by think-aloud, observation, and stimulated recall. Traditional teacher cognition studies mainly focus on individual teachers' beliefs and practices without in-depth study on how individual cognition evolves through and interacts with its social context in which individual teachers participate.

It is the research space above that this thesis seeks to occupy, through an in-depth case study of the beliefs and practices of sixteen New Zealand university tutors who were employed in one of the university's faculties to provide feedback on undergraduates' assignments. In addition to exploring the beliefs and practices of this specific group of tutors, and the factors that influence these, the study aims to contribute to both the theoretical and methodological construction of teacher cognition studies by employing a holistic socio-cultural frame work based on Vygotsky' s key notions of cognition, distributed cognition, and an activity theory approach.

Data were collected chronologically across an academic year by five methods: preliminary survey for bio-data of participants and their general attitudes to giving feedback across the faculty, individual interviews for beliefs on giving feedback, think-aloud sessions on the actual practice of giving feedback, stimulated recall discussions as reflection in action, and focus group discussion as a means of collective reflection of various factors underlying their beliefs and practices. Data were firstly transcribed, stored, and open coded by NVivo8 for preliminary analysis and then analysed manually for deeper understanding of themes. Constant comparisons were made through the whole process of data analysis between data from different participants and between different sources of data.

The findings reveal that there were convergences and divergences among tutors between their beliefs and practices about providing assessment feedback to the written work by undergraduate students. The convergences and divergences were due to the contextual factors in the activity system and tutors' previous experiences. The convergences and divergences of tutors' beliefs resulted in emotional reactions. Tutors' emotion interacts with cognition and actions (ECA interaction). The ECA interaction is affected by contextual factors in the activity system. The contradictions of the activity system constrain tutors' cognition, cause negative emotions, and are often barriers to tutors' work, but also form the potential of cognitive development. Co-operative effort is needed in the wider context of the activity to facilitate tutors' cognitive development, promote positive emotions, and achieve a better outcome for the activity.

It is concluded that a holistic socio-cultural framework of teacher cognition contributes to the understanding of the complexity of teacher cognition. The study is significant for its practical implications for professional practice of assessing disciplinary writing and tutor development; its contribution to the development of teacher cognition and activity theory regarding the interaction between emotion, cognition, and action at both individual and distributed level; and a multi-method approach to teacher cognition studies.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

LIST OF TABLES

LIST OF FIGURES

TRANSCRIPTION CONVENTIONS

This thesis was transcribed by the researcher. The transcripts quoted in the thesis were checked by a fellow researcher. The following conventions abbreviated from Du Bois (2006) were used in the thesis.

Table 1: The transcription conventions abbreviated from Du Bois (2006)

MEANING	SYMBOL	COMMENTS
Participation Speaker attribution	J –	Dash follows initial in CAPS, pseudonyms used
Pause Lag/prosodic lengthening	:	Colon marks slowing of local tempo, segment lengthening
Sequence Overlap / short intrusion	[ ]	
Boundary Tone/Closure Terminative Continuative Truncated intonation unit Appeal	. , — ?	Intonation signalling finality (full stop) Intonation signalling continuation (comma) Aborting projected IU (dash) Combines with final/continuing: ?. ?,
Dysfluency Truncated/cut-off word Other dysfluency	wor– < . . . >	Aborting projected word (en dash) Extended stammer, recast, etc. no message conveyed, time indicated by dots for seconds
Vocalisms Laugh Laughing word Vocalism	@ @you're@kidding (COUGH)	One per pulse or particle of laughter Laugh symbol marks laughter during word Various notations: (SNIFF), (AHEM), etc.
Manner Manner/quality Emphasis Quoting	<MISC> </MISC> you're KIDDing 'go go' he said	Various notations for manner of speaking Capitals for strongly stressed syllables Single speech marks indicate a change of voice
Metatranscription Unintelligible Uncertain Transcriber comment	### #you're #kidding [[continues]]	One symbol per syllable Transcribed words are uncertain Miscellaneous comments: 'continues', etc.
Gesture Gesture	((action))	Various notations: ((action)), ((pointing)), etc.
Grammar Sentence start	Capital initial	Capitalize for beginning of new discourse 'sentence'

Note: Unlike Du Bois (2006) who used dot to indicate pause, slash '/' was used in the transcripts of this thesis. Each slash represented approximately 1 second, e.g. /// indicated a pause of three seconds.

CHAPTER ONE: INTRODUCTION

1.0 Introduction

The present study focuses on the beliefs and practices of university tutors, as they constitute a sub-set of teachers more generally, and they have also tended to be under researched compared with other sources of teachers. It investigates how tutors' cognition interacts with emotion and action at both individual and social levels, using a multi-method approach to data collection. It is a case study in the working context of tutors assessing the written work of undergraduates in a New Zealand university.

This chapter briefly outlines the study by introducing the motivation of the researcher, the research spaces identified in the relevant literature, the research aims and questions, methodological framework, the context of data collection, working definitions of key terms, the significance and limitations of the study, and the structure of the thesis.

1.1 Initial motivation and identity of the researcher

This study of tutor cognition about assessment feedback and their practices when assessing undergraduates' written assignments arises out of my personal interest in this issue because of my professional and academic identity and experience. Professionally, I have about fourteen years of English teaching experience in a university in China. Assessing written work by students of various disciplines was a routine teaching practice for me, yet I was often confused because there was no agreed standard to follow. My interest in this issue increased when I studied for a Master of Education degree at Waikato University. I began to think about teachers' expectations of my written assignments and paid special attention to feedback and evaluation given on my writing. I received feedback from different teachers on my written assignments, most of which was helpful to me. However, I also noticed that feedback for one assignment was usually not applicable to another. Different teachers seemed to have different beliefs about good writing. Therefore, I became curious about the sources of teachers' and tutors' beliefs about giving feedback on students' assignments.

Through a literature review of teacher cognition studies on giving feedback, the following research spaces are identified: first, although teacher cognition has been explored widely, the beliefs and practices about university tutors' professional activities such as evaluating and giving written feedback on students' written work have rarely been explored. Secondly, very few studies on teacher cognition in giving feedback have included data of actual practice collected by think-aloud, observation, and stimulated recall. Thirdly, traditional teacher cognition studies focus mainly on individual teachers' beliefs and practices without in-depth study on how individual cognition evolves through and interacts with the social context and distributed cognition (Salomon, 1997) among the communities of practice (Lave & Wenger, 1991) in which individual teachers participate. Moreover, there is little explanation on how emotion interacts with cognition at either individual or social levels in the activity of assessment.

Therefore, an in-depth study on tutor cognition in giving feedback which goes beyond individual level of analysis will significantly contribute to the existing literature of teacher cognition studies.

1.2 Aims and research questions

In addition to exploring the beliefs and practices of this specific group of tutors, and the factors that influence these, the present case study aims to contribute to both theoretical and methodological construction of teacher cognition studies by employing a holistic socio-cultural framework based on Vygotsky's (1986; 1978) key notions of cognition, distributed cognition, and an activity theory (Engeström, 1999) perspective. The research questions that guide this study are:

--What do subject tutors in the specific context believe about giving written feedback on students' assignments?

--What are their actual practices when giving feedback?

--To what extent do their actual practices converge with or diverge from their beliefs, individually and across departments/disciplines?

--What are the socio-cultural factors that influence tutors' beliefs and practices?

--How do the findings to the questions above add to academic understanding of what constitutes teachers' beliefs, and the possible tension that arises in

putting these beliefs into practice?

1.3 Methodological framework

This study is a qualitative case study within an interpretive paradigm. Its purpose is to explore beliefs and practices in the natural context in which they are expressed and affected. Data collection and interpretation were carried out from the perspective of the socio-cultural model of the study which is mainly based on Vygotsky's (1978) social psychological theories. This perspective is reflected in the following aspects: first, the unit of analysis of the case is an activity of university tutors giving feedback. Second, data are collected in multiple ways from the natural context of these tutors' work in order to provide a thick description of the activity. The activity system of giving feedback at the faculty and the university level also informs the context of data collection. Third, data collection methods are selected and synthesized in different stages of the research process to provide a rich interpretation of the participants' beliefs and practices. The process of data collection and analysis is re-examined through a socio-cultural perspective to fit into the qualitative interpretive paradigm and to fully address the research questions.

1.3.1 Data collection methods

Five methods of data collection were used in this study: a questionnaire with closed and open-ended items, individual interviews, think-aloud and stimulated recall sessions, and focus group discussions. All data (except those from the survey) were audio-recorded, first transcribed, stored, and open coded by NVivo8 for preliminary analysis and then analyzed manually for deeper understanding of themes. Constant comparisons and contrasts were made throughout the whole process of data analysis between data from different participants and between different sources of data.

1.3.2 Context of data collection

Assessing undergraduates' writing is a controversial issue among various schools of thought in applied linguistics which has not been explored in sufficient depth in previous teacher cognition studies. Moreover, existing published studies have often only focused on one of the following aspects of assessing undergraduates' writing: error correction as feedback to students who used English as additional

language, feedback on drafts, assessing writing in disciplines regarding content and formal issues, educational assessment techniques, and teachers' assessment literacy. However, the issues above were inseparable in the context of assessing undergraduates' writing in disciplines in my study.

In this study of teacher cognition on assessing undergraduates' writing, I recruited tutors who were engaged in giving feedback to undergraduate students in a faculty of a New Zealand university. The faculty I chose enrolled both native and international students who had varied language backgrounds and prior knowledge of writing. Students, especially those who enrolled in various papers in Arts and Social Sciences, were assessed mainly by the written work they submitted to fulfil the requirements of various courses.

The faculty in this study is composed of ten departments of Arts and Social sciences together with two research centres (for ethical reasons, the names of the departments are not provided here). Nine of the ten departments have the practice of giving written feedback on undergraduate students' written work. The faculty was chosen as my research domain because of its organisational unity, complexity, and the convergence of a wide range of disciplines within one community of practice. Teaching staff in the nine departments are professors, associate professors, lecturers and senior tutors. Due to the large enrolment in some papers, tutors and sessional assistants are recruited each school semester by the department. The main responsibility for most tutors was running tutorials and marking assignments. Some tutors were only employed to mark assignments and their work was overseen by the lecturers or senior tutors. The number of these tutors varied in different departments in different semesters according to the number of students and the status of finance. These tutors (pseudonyms are used for the tutors who participated in my research) were mainly recruited from students at the faculty at PhD or Master's level, or even third year undergraduates. These part-time tutors, like those full time tutors, varied in their background of study and tutoring experiences.

Three points need to be clarified regarding the identity of tutors before the literature review. Firstly, the subject tutors in my study of assessment activity are, except for two full time tutors, part-time sessional assistants. However, the subject

tutors carry out the same activity of assessment as the senior tutors, lecturers and professors who supervise them. Therefore, the subject tutors are members of the community of teaching practice in the university. Subject tutors are generally in marginal positions in the community of teaching practice, so their professional cognitive development can provide insights into the institutional construction of the community of teaching practice at all levels of seniority.

Secondly, subject tutors are not experienced language teachers. However, in the activity of assessing written work and providing assessment feedback, language is the main tool of mediation in their work and they are unavoidably engaged with language issues in tertiary education. As tutors have to mark written work of both native and non-native English speaker students, they have to concern themselves with both language and rhetoric, which are foci of discussion among applied linguists including those who teach English to non-native English speaker students (hereafter referred to as language teachers) and those who teach English composition to native speakers of English (hereafter referred to as compositionists).

Thirdly, recent empirical studies focusing on assessment and feedback on undergraduate students' written work have revealed that subject teachers (who are in disciplinary areas other than academic writing at university) are cognitively confused by the professional requirements of their practice of assessment feedback (Bailey & Garner, 2010). However, conventional teacher cognition theories are unable to address the complexity of teacher cognition because of the narrow focus on individual teachers' cognition. The present study, through an in-depth investigation of the beliefs and practices of a specific group of university staff (tutors) regarding one aspect of literacy - the assessment of undergraduates' written assignments - not only occupies a specific research space but also moves towards developing a theoretical framework in which the complexity of teacher cognition can be more fully addressed.

It is these three considerations that make this case study a unique contribution to the theoretical development of teacher cognition studies by synthesizing different schools of thought in studies of cognition, teacher cognition, assessment and feedback with practical implications of assessment activity and tutor development. In addition, the multi-method research design of the present study (to be discussed

in Chapter Three) will contribute to research methodology in teacher cognition studies.

1.4 Working definitions of terms used in this study

Students' written work: In the scope of this study, students' written work refers to written work that undergraduate students are required to write in various disciplines of the faculty of the New Zealand university. In the present study, these comprised short summaries, reviews, essays, lab reports as well as responses to short answer questions.

Subject tutor: Subject tutors in this study are teaching assistants who have relatively advanced content knowledge in a specific subject and whose main duty is to help lecturers of various disciplines (except those who teach academic writing) to provide academic learning support to students. They have various responsibilities, such as giving tutorials to students, but the major responsibility is to evaluate students' written work. They may be employed full-time but most of them are employed part-time, but all are regarded as a type of teacher or teaching assistant who carry out professional educational activity. It is important to note that none of the participating tutors had received formal professional training in the work of assessing writing and providing feedback on students' written work.

Belief: Belief in this study is coterminous with cognition. It is a mental system which not only includes tutors' existing knowledge and thoughts resulting from their past experience, but also includes the ongoing thinking and decision-making process during the activity of assessing and providing feedback on students' written work.

Practice: Practice refers to the actions tutors carried out during the activity of assessing and providing feedback on students' written work.

1.5 Significance of the study

The present study is primarily significant in its contribution to the theoretical development of teacher cognition studies by a synthesized socio-cultural approach and to activity theory by the explanation of the interactions between emotion,

cognition, and action (ECA) in the activity of assessment. Secondly, it makes a methodological contribution by applying a multi-method approach to data collection in the context of the participants; particularly in the use of Think-aloud and stimulated recall techniques. Thirdly, it has practical implication for the practice of giving feedback and tutor development.

1.6 Organization of the study

The thesis is composed of six chapters. This first chapter has introduced the theoretical and contextual background of the study and outlined the gap in literature. It then introduced the purpose, research questions, methodology, and contributions of the study, and finally outlined the structure of the study.

Chapter Two is the review of the relevant literature. It summarizes the literature of the studies of teacher cognition and identifies two research spaces in teacher cognition studies: one is the insufficient understanding of tutor cognition of assessing and providing written feedback on undergraduates' written work; the other is the general theoretical tendency of teacher cognition studies to be based on individual teachers, which is too narrow a focus to explore social origins of cognition. It then reviews the studies on feedback on undergraduate students' written work with a focus on the divergences of beliefs among different communities of practice (including those who teach writing and those who teach other subjects) and the mismatch between beliefs and practices. The review of these studies demonstrates the practical need for in-depth study on tutor cognition on giving feedback. It also demonstrates that current studies on teacher cognition on giving feedback also have limits in their sources of data (such as survey and interview) besides the above mentioned research spaces in teacher cognition studies in general. The final part reviews socio-cultural approaches to studies of cognition.

Chapter Three covers the methodology used. It first introduces the interpretive research paradigm applied to this qualitative case study. It then discusses the research methods in this study both at theoretical and practical level with the socio-cultural perspective. At the theoretical level, it explains the selection for this study of a range of data collection methods: survey, interview, think-aloud and stimulated recall sessions, and focus group discussions. The advantages and

disadvantages of each method are discussed. Also discussed is how and why these methods are used in the present study. At the practical level, it discusses some detailed issues of data collection in relation to the ethical issues. The last part of this chapter explains the detailed process of data collection and analysis.

Chapter Four presents the findings. Data collected by different methods are presented in the sequence of beliefs and practices of assessing writing and giving feedback, the sources of the beliefs, and emotional factors of assessment. The study found that the tutors had both convergent and divergent beliefs and practices regarding their standards of written work, feedback, grading, and the use of criteria. The study also found that the tutors' beliefs were not always convergent with their practice. The major divergence between the tutors' beliefs and practices was their believed goal of facilitating learning improvement by feedback, and their actual practice of using feedback to justify the grades they allocated. The divergences and convergences were derived from the tutors' previous experiences and the current context of the activity of assessment. The study found that the tutors had emotional reactions at work. Tutors' emotions interacted with cognition and action in the process of assessing students' written work. The tutors also used strategies to reduce students' potential negative emotional reactions toward assessment feedback.

Chapter Five discusses the findings in relation to the key works reviewed in Chapter Two. It first applies the key principles of Vygotsky's (1978, 1981) theory of the role context plays in the development of an individual's cognition and the role language plays in mediating the process of internalization and externalization. It then analyses the contextual factors with Engeström's (1987) expanded model of activity theory. It argues that the inevitable changes within a system give rise to contradictions in beliefs and practices, which then cause cognitive and emotional reactions. The causal relationship between emotion and cognition is explained by the concept of cognitive dissonance (Festinger, 1957) and cognitive appraisal (Gross, 2002, 2008). It then explains the interactions of emotion, cognition, and action (ECA) at both individual and collective levels in the activity system. It analyses the contradictions within the activity system and argues that these contradictions cause emotional reactions, and affect tutors' work. Therefore, more collaborative efforts are needed in the wider context of the university in that

assessment is related to other activity systems in the university especially in regard to policy-making and management activities. It concludes that the study of tutor cognition of assessing witting is significant in that it reveals the roles that context and emotions play in the tutors' cognition and practices and how a multi-method of data collection can contribute to revealing these factors.

Chapter Six is the final chapter of the thesis. It concludes that tutors' cognition interacts with emotion and action. The ECA interaction is regulated by both self and others in the context of the activity. A holistic socio-cultural framework of teacher cognition and a multi-method approach of data collection contribute to an understanding of the complexity of teacher cognition. It ends with the implications for further research on the professional development of tutors such as those in focus in this study, and for the theoretical and methodological construction of teacher cognition studies.

CHAPTER TWO: LITERATURE REVIEW

2.0 Introduction

Chapter Two reviews the theoretical and methodological issues involved in teacher/tutor cognition studies and demonstrates the complex context of teacher/tutor cognition. It argues for the possibility, indeed the necessity, of making clear connections between studies regarding theory and policy with regard to assessment, feedback, writing-across-disciplines, assessment literacy, teacher cognitive development, and professional training, all of which usually collectively influence teacher cognition in teachers' professional practice, leading to the research space that this thesis seeks to occupy.

This chapter reviews relevant literature relating to beliefs and practices of university tutors as regards the activity of assessing the written assignments of students in various academic disciplines. It firstly reviews the relatively limited number of studies on subject tutor cognition of assessment feedback with an analysis of theoretical and methodological issues within these studies. This is followed by a review of the context of university teacher cognition of assessment feedback. The context here includes theoretical and institutional contexts that influence university teachers, including subject tutors in general. It then reviews conventional teacher cognition theories and empirical studies on teacher cognition studies to demonstrate that conventional teacher cognition theories and research methods are insufficient to address the complexity of university teacher cognition. The following part reviews socio-cultural theories that can be used as a synthesized framework to address the existing issues, both theoretically and methodologically, in teacher cognition studies in general and tutor cognition of assessment feedback in particular. The last part of the literature review is the summary of research spaces which this study will seek to occupy.

2.1 Assessment feedback: Subject tutors' beliefs and practices

This section reviews studies of the identities of subject tutors and their professional background, studies on subject tutors' practices of assessment feedback and the theoretical and methodological issues which arise in these studies. It intends to demonstrate the particularity of subject tutors as a sub-

category of university teachers and the limitation of studies of subject tutor cognition on assessment feedback.

2.1.1 Identities and professional background of subject tutors

The role, status and teaching activity of subject tutors vary in different universities. They can be loosely defined as teachers who carry out teaching activities ranging from running tutorials to marking assignments and supporting learning in various disciplines. Some tutors are temporary, working as part-time sessional assistants, while others have full time permanent positions. In the hierarchical structure of a university, tutors are usually at a place peripheral to the teaching community but often have the most contact with students. These tutors may have various educational backgrounds but usually have subject-specific knowledge in the discipline area in which they work and often have no teacher training background. Smith and Bath (2004) note that training programmes for new tutors or sessional staff in universities in Australia and UK are underdeveloped compared with those in North America. The same issue has also been found in New Zealand (Sutherland, 2009).

The lack of institutional support for part-time tutors and even for full time teachers has been explored by Knight, Tait, and Yorke (2006) in two studies: the first study was based on survey response from 2401 part-time tutors in the Open University in UK, 92 electronic interviews of survey respondents, and 43 follow-up telephone interviews among the 92 interviewees; the second study was based on survey response of 284 full time teachers. The finding of the studies was that learning-by-doing was the main type of professional development among both part-time tutors and full time teachers. This finding is similar to that of Sutherland's (2009) study on university tutors in New Zealand. Sutherland interviewed twelve senior undergraduates working as tutors across six university departments about their practice in the teaching community and the support they needed. The study revealed the benefit tutors brought to the teaching community, such as the insights they had into the needs and level of understanding of students, and the easy access to help they provided to students. Tutors also benefited from their participation in the teaching community regarding their communication skills and relationship with lecturers. However, the study also raised the issue of the lack of teaching experience and training of these tutors, and concluded that

tutors needed "structured and systematic support" (p. 159). Therefore it is necessary to discover how tutor cognition develops at work and what professional support is needed for them to work effectively.

2.1.2 Studies of subject tutors' practices of assessment feedback

Conventionally, studies on feedback may refer to either feedback on drafts or assessment feedback. Teachers who provided feedback may include tutors, lecturers, or professors. Very few studies have made clear distinctions between the nature of feedback and/or the identities of those who provided feedback.

The limited number of studies of tutors' beliefs and practices has focused on, and provided understanding of, two issues: the divergences and convergences among tutors regarding their beliefs and practices of feedback, and constraints on tutors' practices.

Divergences and convergences of beliefs and practices among tutors who provide assessment feedback have been revealed in the study conducted by Ivanic, Clark, and Rimmershaw (2000). They compared the feedback on nine pieces of writing by five subject tutors of social sciences in UK and four EAP (English for Academic Purpose) tutors in two UK and South African universities. They found that subject tutors varied greatly regarding whether to give feedback and how much feedback to provide. Some tutors red-marked the errors on students' writing, while others used a pencil. Some wrote on the margin while others attached a separate sheet. However, the following five aspects were common among the subject tutors: students' writing was regarded as final product rather than work in progress and no tutor responded to drafts; tutors read assignments for the purpose of grading and their feedback was mainly used to justify the grade; all subject tutors pointed out more negative than positive aspects and evaluated assignments against the expected answer; few tutors proofread students' work; almost no subject tutors' feedback indicated an engagement in on-going dialogue with students. The study has its shortcomings in that it neither described the context of teaching in detail, nor did it reveal the influence of contextual factors on participants' behaviours. Also, there was no description of how data were collected and analysed, and apparently, no data were collected by observing tutors' actual practices of giving assessment feedback. However, the research has two

12

implications: first, tutors tend to give feedback according to their own values and beliefs about writing; secondly, feedback may have a negative effect on students' confidence.

Other studies have also found divergences and convergences between tutors' beliefs and practices. Orsmond and Merry (2011) studied the link between tutors and students' understanding of feedback in a British university. Data were collected from 19 students and six tutors by interview and document analysis of written feedback. Orsmond and Merry found that the tutors tended to use corrective feedback and provided praise on good points, which practice matched these tutors' expressed beliefs. They also found that tutors believed feedback should be given on how to improve the written work; however, this belief was not put into practice in their written feedback.

The other issue, constraints on effective practices of tutors, has been explored by two recent studies (Bailey & Garner, 2010; Lilly, Richter, & Rivera-Macias, 2010). Lilly et al. compared tutors' and students' opinions and experiences of assessment feedback by documentary analysis of samples of good feedback, as well as focus groups and workshops among teachers and students across a UK university. They found tutors across faculties had common issues that constrained their practices, including personality, lack of time, poor communication and understanding of feedback practices, the modular pattern of courses, and organizational issues. The findings of Lilly et al.'s study concur with those of Bailey and Garner, except that the latter provided more information on social-institutional constraints. Both studies found that tutors needed professional guidance.

2.1.3 Summary

Subject tutors are a group of university teachers whose cognition of assessment feedback has not received enough attention in teacher cognition studies. Current studies have revealed there are convergences and divergences among tutors regarding their beliefs and practice. Tutors face contextual constraints and need professional support in their work. However, current studies are predominantly dependent on survey, interview, and/or document analysis rather than observational data; neither have current studies employed a theoretical framework

that can fully explain the nature of tutor cognition of assessment. The present study will address these two key issues.

2.2 Contexts of subject tutor cognition of assessment feedback

Tutor cognition of assessment is influenced by both theoretical perspectives and institutional policies. Theoretical perspectives in this study refer to schools of thoughts in discourse communities (Swales, 1990), mainly of applied linguists and compositionists, that focus on assessment of and feedback on undergraduates' written work. The institutional policies include the requirements of teaching practice, professional ethics and professional training that relate to assessment within institutions. The following section will first review the theoretical perspectives and then the institutional contexts of assessment and feedback.

2.2.1 Theoretical perspectives on assessment and feedback

Generally speaking, studies of assessing undergraduates' writing and providing written feedback are carried out in the area of applied linguistics and composition studies. The two groups of studies overlap and to some extent inform each other while having different foci. Feedback on formal errors in the drafts written by students for whom English is an additional language is mostly addressed in the area of applied linguistics, especially in studies of second language writing; whereas feedback on written work of native speaker students in various subjects is mainly discussed in composition studies. Both areas intend to inform teaching in disciplinary areas of good feedback and assessment guidelines. However, the issues discussed in the following sub-sections cause confusion among teachers, especially tutors, in their beliefs and practices of providing assessment feedback. These theoretical issues include multi-purposes, multi-foci, strategies, and criteria of assessment feedback.

2.2.1.1 Purposes of assessment and feedback

According to Price, Handley, Millar, and O'Donovan (2010), feedback on students' written work has not been clearly defined, regardless of different schools of thought on written feedback in applied linguistics, composition, and educational evaluation.

In an educational context, feedback is often used as a goal-oriented pedagogical tool facilitating students' learning improvement. For example, Keh (1990) stated that feedback should be goal-oriented in relation to the current progress and future improvement. Nicol and Macfarlane-Dick (2006) listed seven principles of good feedback: clarification of goals to both students and teachers; providing the right information of current learning progress to enhance learning improvement; providing students with information about learning; encouraging self-regulation, self-esteem and motivation; opening dialogue between teachers and students; and informing teaching. Similarly, Hattie and Timperley (2007) suggested that feedback should address the goals and the current progress toward the goals. They also pointed out that feedback should include information on how students could make improvements to achieve their goals.

However, feedback serves more than one purpose when it is used in assessment. Theoretically, formative assessment aims for improvement of learning and is carried out during the learning process, while summative assessment is used at the end of a course as an evaluation of learning and teaching (Biggs & Tang, 2007). The main purpose of formative assessment is to assist students to improve their work via feedback (Black & Wiliam, 2009; Sadler, 1998). The main purpose of summative assessment is believed to provide valid and reliable measurement of learning outcome (Cherry & Meyer, 1993; Moore, O' Neill, & Huot, 2009). However, in practice, assessment is usually both formative and summative in most courses in various disciplines (Lea & Street, 2000), and thus the dual purpose presents some problems (Biggs & Tang) because students tend to focus on marks or grades rather than the written comments (Butler, 1988; Carless, 2006; McGee, 1999; Mutch, 2003). It is argued that assessment feedback that focuses on measuring students' achievement rather than enhancing improvement is ineffective (Hattie & Timperley, 2007). Some recent studies argue that summative feedback should also play a formative role by suggesting to students how to improve (Hounsell, McCune, Hounsell, & Litjens, 2008; Lilly, Richter, & Rivera-Macias, 2010).

In addition, neither feedback nor assessment aims only at the improvement of learning and/or the evaluation of current achievement because, according to Joughin (2009), in addition to these goals, assessment is intended to maintain

disciplinary and professional standards. Cooksey, Freebody, and Wyatt-Smith (2007) argue that the goal of feedback is unavoidably influenced by the goal of assessment which involves the "differing interests, needs, expectations, and preferred discourses" (p. 402) of various stakeholders which are selectively applied by teachers in the process of assessment. Therefore, the literature of assessment theory may seem confusing for teachers because of the dual function of feedback for improvement and measurement (Taras, 2006). It seems ambitious to attempt to achieve multiple goals for teachers, given the ambiguous nature of the disciplinary standards, and the insufficient strategy in literature on how to support learning via assessment feedback. This is especially so for subject tutors, because of their marginal status in both disciplinary and professional communities,

To sum up, there is increasing consensus on the formative role that assessment feedback should play, and a growing realisation in the literature that feedback serves multiple purposes. However, there have been few strategies suggested for how the formative and summative function of assessment feedback can help students improve their writing. This issue is related to the range of institutional constraints to effective assessment that will be reviewed in Section 2.2.2.

2.2.1.2 Foci of assessment feedback on writing in academic disciplines

The focus of assessment feedback relates to the disciplinary values and goals of education. Compositionists believe that writing is "central to students' success when they enter college, during postsecondary education, and into careers" (Adler-Kassner & Harrington, 2010, p. 76). The problem in discussion is how much value should be given to the form of writing to achieve the goal of disciplinary education (Carter, 2007) and how much of the feedback provided should be given on the formal features of language and writing as against content, especially in contexts where students in various disciplines are composed of both native and non-native speakers of English. Moreover, there are issues about which formal aspects can be and should be commented on considering that academic written work includes both discipline-specific genres (e.g., lab reports, expository essays) and non-discipline specific forms (e.g., short answers) (Melzer, 2009).

There is debate in the applied linguistics literature on what aspects of writing should be the foci of feedback. According to Keh (1990), feedback should focus

on higher-level concerns of writing such as text organization. However, feedback on lower level concerns, such as grammatical errors, has been the focus of discussion among those applied linguists who are primarily concerned with non-native speaker students' writing. A large group of studies have argued (or suggested) that error correction is effective (Bitchener, 2008; Chandler, 2003; Fathman & Whalley, 1990; Ferris, 1999, 2003, 2004, 2006), whereas others (e.g. Truscott, 1996, 1999, 2007) have opposite opinions. The focus of argument between Truscott (1996) and Ferris (1999) is whether teachers have sufficient knowledge, ability, time and effort to provide appropriate feedback on grammatical errors according to each learner's development process.

A number of empirical studies on subject teachers' assessment feedback have found that the grammatical errors are the focus of feedback provided by subject teachers; little feedback is given on disciplinary features of writing. For example, Lea and Street (2000) found that subject teachers in two UK universities often could not describe explicitly the features of good argument in an assignment. This finding is in agreement with that of Stern and Solomon's (2006) study on feedback provided by instructors on 598 assessed written assignments collected from 30 departments in a university in US. Moreover, Stern and Solomon found that the focus of feedback in formative assessment was lexical level errors; little feedback was provided on organization of the writing. Perhaps the focus on lower level concerns is due to the assumption among teachers of various disciplines that "faculty should not be asked to articulate or teach the communication conventions of their disciplines" (Barlow, Liparulo, & Reynolds, 2007, p. 54).

In sum, the issues remain theoretically debated regarding how much feedback should be provided on writing and to what extent formal aspects of writing, rather than the content of the written assignment, should be the focus of feedback. In contrast to the focus of feedback suggested in the literature, the findings of empirical studies have revealed that assessment feedback usually focuses on lower level concerns. These discussions are relevant to my study on the subject tutors who assessed written work of both native and non-native undergraduates in various disciplines.

2.2.1.3 Strategies to improve the effectiveness of assessment feedback

Integrative feedback is the most widely-advocated strategy of good feedback practice which is widely accepted by both compositionists (Faigley & Witte, 1981; Sitko, 1993) and writing teachers of second language speakers (Broad, 2003; Cohen & Cavalcanti, 1990; Dheram, 1995; Fathman & Whalley, 1990; Leki, 1990). The key feature of this model is to allow students the opportunity to submit drafts, respond to feedback and make revisions. Portfolio assessment has been advocated as an effective strategy to record the progress of students' writing and to provide students with the opportunity to reflect on feedback over time (Weigle, 2007).

Another strategy which has been suggested is "assessment dialogues" (Carless, 2006, p. 230) between tutors and students with mutual understanding of the content and roles of criteria, the expectations of tutors and students on the assessment and opportunities for improvement. Nicol (2010) also points out the need to change the conventional feedback from monologue to dialogue between teachers and students. He further suggests that the possible increase of workload could be shared by combining teacher feedback with peer feedback.

Studies of disciplinary writing have also demonstrated the effectiveness of feedback on preliminary drafts. Fisher, Cavanagh, and Bowles (2011) carried out a multi-method study on the effects of oral and written feedback on drafts of literature reviews during the first semester of a first year business course in an Australian university. Positive results were found in the improvement of students' marks and their understanding of teachers' expectations. Similarly, Duijinhouwer, Prins, and Stokking (2010) collected survey data among students who received progressive information in the feedback of written work in a psychology course in a university in the Netherlands. Their data demonstrated the effectiveness of progressive feedback on the improvement of writing. However, both studies were carried out in a controlled context for the purpose of hypothesis testing rather than in a natural working context. Studies carried out in natural contexts indicate that feedback on drafts is not a common practice adopted by teachers/tutors in disciplinary writing (Ivanic et al., 2000; Lea & Street, 2000). In an Australian context, Brinkworth, McCann, Matthews, and Nordström (2009) surveyed both first and second year undergraduates and teachers (including lecturers and tutors)

in humanities and science majors regarding their expectations and experiences of teaching and learning, including written feedback. According to the survey result, feedback was rarely given on drafts.

To sum up, these strongly recommended strategies or models are not widely adopted by subject teachers because assessment feedback on written work is constrained by various contextual factors to which attention will be turned in section 2.3.2.

2.2.1.4 Reliability of assessment

Reliability refers to whether individual markers of one task can keep consistency of assessment within a cohort of students and also whether different markers assess the same assignments similarly (Jonsson & Svingby, 2007; Reddy & Andrade, 2010). Reliability is a concern of this study in that tutors who mark the same written assignments are expected to keep consistency with each other and with their lecturers.

It has been found that the use of pre-determined criteria enhances consistency, especially when marker training is provided (Jonsson & Svingby, 2007; Reddy & Andrade , 2010). However, inconsistency exists across markers who assess the same piece of writing due to their different backgrounds (Hamp-Lyons, 1996; Pula & Huot, 1993; Vaughan, 1991; Weigle, 1999). For example, Vaughan (1991) studied marker variables by asking nine trained experienced markers (who were language teachers) to talk into a tape recorder while they were holistically marking six essays written by former undergraduate students. Divergent results of marking were found due to individual marking styles and the fact that essays often fall between scales. The study also found that when marking groups of assignments, markers often formed their judgements by making comparison between essays.

Barkaoui (2007) explored the marking perspectives and processes of assessment of four language teachers at a university in Tunisia by a mixed-method approach including questionnaire and two sessions of think-aloud (TA) during which the markers were using holistic and analytical marking scales to assess four essays written by language learners. Training was provided before participants were

asked to think aloud. There were discussions on rating scales before TA and follow-up interviews immediately after TA. Like Vaughan (1991), Barkaoui (2007) found that markers had different marking processes. Moreover, Barkaoui found that the same criteria for writing assessment were interpreted differently by different markers. Comparison of essays and self-generated criteria were used more often when markers used holistic criteria. When they were marking with analytical scales, more interaction was found between markers and marking scales.

The studies on marker variables revealed that inconsistency exists even among experienced markers; the common strategies to avoid inconsistency included training of markers, using criteria, and making comparisons.

Having reviewed the theoretical perspectives of the purpose, foci, strategy and reliability of assessment, the following sub-section will review the influence of policies on the activity of assessment.

2.2.2 Institutional polices

According to Baker (2010), socio-cultural issues of assessment and feedback have long been noticed yet not fully explored, and little is known about how individuals interact with the context in their assessment practice. It can be argued that the broadest socio-cultural contexts of university teacher/tutor cognition are at international and national levels, for example through professional associations, and conferences. However, the direct context of teacher cognition is at the institutional level.

2.2.2.1 Institutional requirements of assessment

The institutional factors that influence teacher cognition of assessment include variation in students' writing background (Hamp-Lyons, 1991; North, 2005; Sakyi, 2000); the number of students (Goldstein, 2005); various institutional requirements such as structured feedback forms (Bailey & Garner, 2010); the modular patterns of courses which tend to have negative effects on the effectiveness of feedback (Price et al., 2010); different expectations of writing in different courses (Lea & Street, 2000; Lilly et al., 2010); and application of

different criteria in the assessment (Barkaoui, 2007; Becker, 1991). Among all these factors, institutional policies for assessment have been found as major constraints to effective assessment practice.

As has been reviewed in previous sections, the theoretical motivation for assessment and feedback is primarily assumed to be the improvement of learning. However, a review on assessment literature by Price, Carroll, O' Donovan, and Rust (2011) reveals that an institutional policy for assessment is often simplified to comply with the institutional requirements; some institutional requirements of assessment practice - such as using feedback sheets with tick boxes which are assumed make for 'objective' feedback - lead to student dissatisfaction. The contradiction between institutional policies of measurement and the claimed aim of pedagogy of assessment for learning improvement has been found by three recent empirical studies on assessment feedback (Bailey & Garner, 2010; Meyer, Davidson, McKenzie, Rees, Anderson, Fetcher & Johnston, 2010; Price et al., 2010). Bailey and Garner explored contextual factors such as large class size, heavy workloads, and general policies of giving timely assessment which should be both formative and summative. They collected data by interviews with 48 teachers across departments in a British university. Their study revealed that teachers were often in a feedback dilemma by being expected to help students improve via their feedback and at the same time to satisfy institutional requirements. They felt a "lack of ownership" (p. 196) of their own feedback because they had to "adjust their language to meet the perceived needs of individual students, circumventing the limitations of forms and official standards" (p. 196). In order to meet the various requirements, teachers applied various strategies when giving feedback, which might result in inconsistency in their practice, which in turn confused students. Price et al. (2010) investigated the engagement and effectiveness of feedback across students and teachers in business schools in three UK universities. The data, mainly based on interviews of 20 teachers and 15 students, demonstrated that teachers were confused by the various and conflicting purposes of feedback; the effect of feedback was limited by the modular pattern of courses, lack of time, methods of feedback, and insufficient dialogue between teachers and students. They concluded that it was almost impossible to measure the effectiveness of feedback. The issue of conflicting purposes of assessment has also been found by Meyer et al. (2010),

21

who collected data by large scale survey and document analysis on beliefs about assessment among teachers and students in four New Zealand universities. Like the findings of the studies in the UK, the study found that guidelines and policies for assessment focused on general procedural requirements and the purpose of assessment was mainly the measurement of outcomes rather than improvement of learning; the contradiction between claimed purposes and the practice causing confusion among teachers and students.

To sum up, the institutional requirements or policies of assessment are generally not based on pedagogical principles of assessment. Teachers are in a professional dilemma in assessment practice and have often to sacrifice pedagogical principles to institutional requirements. Therefore, it is necessary to have further in-depth exploration of teachers' beliefs and real practices of assessment in New Zealand universities as elsewhere.

2.2.2.2 Professional ethics and professional training for assessment

Ethics of language assessment refers to the standard of professional assessment practice which should be technically reliable and valid, socially fair, and pedagogically instructive (Hamp-Lyons, 1997). Ethics of language assessment covers political, social, technical, and individual aspects of assessment practice (McNamara & Roever, 2006). It seems impossible for teachers to meet the standard requirement of ethical assessment without professional support and appropriate literacy education. The lack of assessment literacy among university teachers has been investigated in a North American background by two studies (DeLuca & Klinger, 2010; Volante & Fazio, 2007). Based on data collected by surveys, both of these studies found that pre-service university teachers did not have sufficient knowledge of assessment. DeLuca and Klinger identify the need for the assessment literacy integrating "practice, theory, and philosophy" of and for learning (p. 424).

In sum, there is a lack of institutional support for subject teachers regarding assessment ethics and standard assessment practice. There is also a lack of relevant studies outside USA. Therefore, one of the significant themes of the present study is the need to explore the relation between the institutional context and tutors' existing knowledge of assessment and difficulties they meet at work, so as to understand what kind of ethical and professional support tutors need.

2.2.3 Summary

This section has reviewed the context of subject tutors' cognition of assessment including different schools of thought on assessment and institutional policies. It has indicated that subject teachers are insufficiently supported both theoretically and institutionally. There is a need for research into the impact of contextual issues on tutors' assessment of undergraduate students' writing. In addition, it is necessary to explore the construct of teacher cognition, to which attention will now turn.

2.3 Teacher cognition theories and studies

This review now considers conventional teacher cognition theories and issues of studies in relation to teacher cognition of assessment. It firstly intends to demonstrate that conventional teacher cognition theories focus on the relation between individual teachers' beliefs and practices without framing the context of teacher cognition and explaining how context interacts with beliefs and practice. Common issues of teacher cognition studies reviewed in this section are based on four comprehensive reviews: Clark and Peterson (1986), Fang (1996), Kane, Sandretto and Heath (2002), and Borg (2006). It then intends to demonstrate that existing studies into teacher cognition of assessment have similar theoretical and methodological issues to those in teacher cognition studies in general.

2.3.1 Teacher cognition theories

Teacher cognition, according to Borg (2009), is about "what teachers think, know, and believe" (p. 163). It includes the mental process which has been described in various terminologies such as "*knowledge* (and its sub-type), *beliefs*, *attitudes, conceptions, theories, assumptions, principles, thinking and decision-making*" (Borg, 2006, p. 272, italics in original). It is "personally defined, often tacit, systematic, and dynamic" (p. 272). Three main branches of studies on teacher cognition have been found in literature: content, sources, and context of teacher cognition.

The main content of teacher cognition, according to early teacher cognition studies, is knowledge and belief. It is believed that teacher cognition is composed of explicit and implicit knowledge and beliefs (Clark & Peterson, 1986). Shulman (1986) divided the content of teacher knowledge into knowledge of disciplinary

content, pedagogy, learners, context, tools, and goals of teaching. This notion has been widely quoted because it specifies three key aspects of teacher knowledge: discipline, pedagogical application, and the application of the knowledge in the specific teaching context.

Studies on sources of cognition have found that teachers' beliefs emerge from their practices. One significant contribution was made by Woods (1996), who developed a cognitive model for the study of teachers' beliefs, assumptions and knowledge (BAK). According to Woods, the planning of classroom events depends on teachers' belief systems and background knowledge structures. Pedagogical actions are carried out with the intention of operationalizing the plans. What happens in practice adds to the teachers' understanding/interpretation. The three elements interact with each other and form a coherent psychological system of teacher cognition (Li, 2009). It is also found that teachers who are engaged in the same activities may hold different beliefs (Breen, Hird, Milton, Thwaite, & Oliver, 2001; Clark & Peterson, 1986) but they also share some common beliefs (Breen et al., 2001). The articulation of beliefs (Richards, 1996) and reflection on practice (Borg, 2006; Schon, 1991) tend to contribute to cognitive development.

Other sources of cognition include teachers' previous observation of teaching (Borg, 2006; Johnson, 1994; Lortie, 1975; Powell, 2002) and perhaps teacher training (Borg, 2006; Cabaroglu & Roberts, 2000; Freeman, 1993; Richards, 1996). However, pedagogical knowledge for university teachers of various disciplines mainly comes from their previous observation (as learners) of their own teachers and their subsequent professional practice rather than any systematic professional development (Boice, 1992; Dunkin, 2002; Lortie, 1975).

The third branch of study, the context of cognition, has been attracting increasing attention. An emphasis on the relation between context and knowledge can be traced back to the notion of situated knowledge proposed by Lave (1988) and Lave and Wenger (1991). The main argument of situated knowledge is that professional knowledge is always embedded and evolves in the context of teaching practice. It has been found that teachers' practices may diverge from their beliefs due to the influence of contextual factors (Basturkmen, Loewen, & Ellis, 2004; Borg, 1998; Burns, 1996; Crookes & Arakaki, 1999; Lee, 2009).

However, the notion of context varies in different studies. A conventional perspective of context in language-related teacher cognition studies is inside or around the classroom where teaching activities are carried out (e.g. Woods, 1996; Borg, 2006). Freeman (2002), in his review of research on language teacher education in North American backgrounds, identifies a changing perspective of context from classroom to school. In a recent study by Cross (2010), context is extended to the "broader social, cultural, historical, and political genesis of the activity" (p. 447).

In sum, current teacher cognition theories have considered the content of teacher cognition, the relation between cognition and practices, and the influences of context on cognition. The evidence of contextual influence on teacher cognition strongly supports the view that cognition is socially constructed. Therefore the context of teacher cognition needs to be carefully defined in relation to the theoretical (both academic and professional), social, political, organizational, individual and collective aspects of cognition. Moreover, there is a need to explain how contextual factors interact with individual cognition in practice. Context is a major concern in this study not only because there is a lack of in-depth research into teacher and tutor cognition in the New Zealand university context, but also – and more importantly - because context is the key to understanding their beliefs and practices. The present study embraces and investigates the broader context of teacher cognition as discussed by Cross (2010).

2.3.2 Issues of teacher cognition studies

There are two common issues in studies of teacher cognition: focus of the study and methodology. Both issues have been noted in four relevant comprehensive historical reviews: Clark and Peterson's (1986) and Fang's (1996) reviews of teacher cognition in general; Kane et al.'s (2002) review of teacher cognition studies at tertiary level and Borg's (2006) review of language teacher cognition studies. This section will review the two common issues of teacher cognition studies, firstly in general, and then with specific focus on the studies of teacher cognition of assessment.

2.3.2.1 Focus of teacher cognition studies

The common foci found in the four reviews were teacher beliefs and practices and the relationship between them. Clark and Peterson (1986) summarized three categories of studies on teachers' thought process in relation to actions of teaching: planning, decision-making, theories and beliefs. Four issues of teacher cognition study identified by Clark and Peterson (1986) were a lack of studies other than those involving primary school teachers; a lack of relationship between thinking and action in a real class context; a lack of research on novice teachers and the evolution of their thoughts; and a lack of description of "tasks and teaching situations that call for thoughtful teaching" (p. 292).

The second issue mentioned by Clark and Peterson (1986) was also addressed in studies reviewed by Fang (1996). These studies reported contradictory findings of consistent or inconsistent relationship between beliefs and practices. The major reason for inconsistency was the complex contextual factors of the activity of teaching. Fang (1996) identifies gaps in the reviewed studies regarding the findings of study: a lack of understanding of how beliefs are applied to classroom context; a need for further research on the issue of consistency between belief and practice; a lack of understanding of teachers' beliefs about specific aspects of a subject area; a need for studies on teacher cognition in incorporating literacy skills into specific content areas; and a lack of study on teacher cognition at tertiary level.

The need for studies on university teacher beliefs was addressed by Kane et al.'s (2002) review of fifty studies on university teachers' cognition. They found that the research questions in these studies were mainly about teachers' reported beliefs about teaching practice in general (Kember & Kwan, 2000; Samuelowicz & Bain, 2002; Trigwell & Prosser, 1996) and the connection between their beliefs and practices (Hativa, Barak, & Simhi, 2001; Martin, Munby, & Hutchinson, 2000; Quinlan, 1999). Kane et al. (2002) suggest that future research should make clear how teachers relate their beliefs to teaching practice at tertiary level, what is unique about university teacher cognition, and how university teachers develop their theory of teaching.

The contextual issue mentioned by Fang (1996) was emphasized in Borg's (2006) review of 180 studies of language teachers' cognition in which he also identifies methodological issues in the study of language teacher cognition.

These reviews indicate some common concerns and historical development in teacher cognition studies such as the relation between cognition and practice, and the influence of context in teacher cognition. However, as pointed out by Kane et al. (2002), these studies have paid much less attention to teachers at tertiary level, especially regarding subject teachers' cognition of assessing writing. This is precisely the focus of the present study.

The very limited number of studies has revealed that there are divergences and convergences among subject teachers regarding their beliefs about and practice of assessment feedback in the same university. Orrell (2006) explored sixteen experienced subject teachers' beliefs about and their actual practices of assessment and feedback. The study found that experienced teachers had divergent practices regarding the length and communicative style of feedback due to the different identities the teachers attributed to themselves such as co-learners, or experts. Moreover, when providing assessment feedback, little convergence was found between teachers' beliefs and their actual practices. Divergences between feedback beliefs and practices have also been found within the same discipline or discourse community. Read, Francis, and Robson (2005) analysed written feedback on undergraduates' written work provided by 50 historians across 24 universities. Large variation was found among the comments and grades on two sample essays. Similarly, divergences between feedback beliefs and practices have also been found among teachers of writing (Cohen & Cavalcanti, 1990; Diab, 2005). Therefore, it seems a common issue that there are convergences and divergences between teachers' beliefs and practices within disciplines, and within and across universities. However, few studies have provided an in-depth explanation on the sources of such divergences which is a major focus of the present study.

Furthermore, emotion, as a factor that affects cognition and practice, has only been explored in earlier studies (e.g. Cooper & Baron, 1977) focusing on students' emotional and cognitive change influenced by teachers' practice. The relationship

27

between emotion and cognition of university teachers has not been investigated in depth in recent studies. As far as I can identify, no research has been done on the interaction between emotion and cognition of subject tutors in assessment activity. The present study will investigate the role of emotion in university tutors' cognition and practice and thus contribute to academic understanding of the relationship between emotion and cognition of assessment activity at tertiary level.

2.3.2.2 Methodological issues

Most current investigations into teacher cognition rely heavily on self-report data from surveys and interviews, which are insufficient to reveal teachers' cognitive process and the impact of contextual issues on their practices. In studies of teachers' beliefs about and practices of assessment feedback, data collection methods are mainly survey, interview, and document analysis (e.g. Bailey & Garner, 2010; Ivanic et al., 2000; Lilly et al., 2010; Orsmond & Merry, 2010; Tang & Harrison, 2010).

Only a few of the studies on language teachers' beliefs about feedback have included think-aloud data (Cohen & Cavalcanti, 1990; Diab, 2005). However, there is a lack of information in these studies on how think-aloud was used to collect data. In some other studies - for example on markers' thinking in the reading process (e.g. Barkaoui, 2007; Crisp, 2008; Vaughan, 1991) - more information can be found on how think-aloud data are obtained. However, these studies have three common limitations. Firstly, the work to be assessed was selected and provided by researchers and the rating process was conducted for the purpose of research rather than as part of the actual institutional activity of assessment. Secondly, there was insufficient explanation on why and whether (Barkaoui, 2007; Crisp, 2008), or not (Vaughan, 1991), training or practice in think- aloud was provided. Thirdly, the think-aloud data were all audio-recorded: there was a lack of observation data on the marking process.

To sum up, studies regarding university teachers' beliefs about and practices of assessing writing have found that there are divergences between teachers' beliefs and practices both within and across universities. However, there is a need for an in-depth explanation of the reasons for the divergences. Most studies in this area have collected data by survey and interview. Think–aloud has not been used

among subject tutors who provide assessment feedback in their actual working contexts. There is a lack of observation data on the process of providing assessment feedback. Therefore the present study occupies a research space for methodological as well as theoretical development in naturalistic studies of university teacher cognition of assessment.

2.3.2.3 Emotion and cognition

Emotion is a contextually situated response to a specific event, and occurs after cognitive appraisal of the events in relation to goals (Lazarus, 1991). Festinger (1957) explains the relation between cognition and emotion by the theory of cognitive dissonance, according to which cognitive dissonance happens when there is inconsistency or contradiction between beliefs and reality. The contradictory beliefs cause negative emotional reactions, such as uncomfortable feelings, which motivate individuals to make efforts to reduce the dissonance.

The interaction between emotion and cognition has social origins. It is argued that people experience social emotions when they interact with each other in an activity to achieve social goals (Berscheid, 1987; Ellis & Harper, 1975; Oatley, 1992; Simon, 1967). Social emotions include social-evaluative and social-relational emotions that refer to how people feel about each other and their relationship (Leary, 2000). Positive social emotions, such as happiness, come from people's positive evaluations and relations with one another (Leary). Positive and negative social emotions are stimulated by, and in turn, regulate social activities (Lazarus, 1991; Zhu & Thagard, 2002). Like cognition, emotions can also be collective or distributed among individuals (Roth, 2007).

Currently, the influence of emotion has been explored in very few studies on students' cognitive development in language learning (Imai, 2010; Mahn & John-Steiner, 2002). There is research space in theoretical explanation between cognition and emotion (Humphrey, Curran, Morris, Farrell, Kevin, & Woods, 2007). Emotion has received even less attention in teacher cognition studies. Only a very limited number of studies have discussed teacher emotion in narrative analysis of teacher cognition and practice (DiPardo & Potter, 2003). Sutton and Wheatley (2003) reviewed the limited number of studies on teacher emotion in various contexts and noted the "teachers' emotions may influence their own

cognition, motivation, and behavior, as well as their students' cognition, motivation, and behavior" (p. 329). However, this study did not explain the issue of teachers' emotions in further depth.

The relation between emotion and cognition is particularly important in the present study not only because tutor cognition is inseparable from emotion but also because the tutors' assessment feedback tends to generate emotional reactions in students. The emotional issue of feedback was firstly addressed by Jacobs (1974), who suggested that negative emotion on receiving feedback could be reduced by providing feedback on positive aspects before negative aspects. Emotional factors have occasionally been mentioned in feedback studies regarding students' reaction to feedback (Race, 1995). It has been found that while praise might generate positive emotional response (Beason, 1993; McGee, 1999), such positive feedback might not actually encourage learning improvement (Beason, 1993; Leki, 1990). On the other hand, negative feedback can cause negative emotions such as disappointment and shame (Trope, Ferguson, & Raghunatahan, 2001),which is likely to harm students' confidence and motivation (Ferris, 1995; Ivanic et al., 2000; James, 2000), attitudes (Storms & Sheingold, 1999), and affect their self-esteem (Ivanic et al., 2000). Consequently, students may not engage with feedback when it is negative (Winter, Neal, & Waner, 1996). Recent studies have found that assessment usually generates negative emotions. For example, Kvale (2007) argues that the dominant purposes of assessment are "for selection and for discipline" (p. 62), which cause anxiety. Falchikov and Boud (2007), by analyzing autobiographical accounts of adult students' assessment experiences, found that students' assessment experiences were in most cases negative. They conclude that feedback, if interpreted as inappropriate, can cause a negative emotional response in students.

Recent studies have also demonstrated that the emotional aspect of assessment and feedback is bi-directional and interactive between teachers and students. One kind of emotion is teachers' empathy towards students in assessment (Hawe, 2003). According to Värlander (2008), empathy for students has a positive effect in that it can reduce the negative emotion caused by the power relationship between teachers and students; it may be easier for tutors than lecturers to show

empathy to students because of their relatively low position in the teaching community.

Another emotional aspect of assessment and feedback is trust. Lee and Schallert (2008) examined interactive factors between teachers and students during the feedback-revision cycle of a composition course in a Korean university. Data were collected by interview, observation and analysis of student drafts and feedback from a teacher and her students. The study revealed that trust between teachers and students played an important role in the effectiveness of feedback. Carless (2009) examined the role of trust in assessment practice in an English department in Hong Kong. He found that distrust could be distributed between the management staff and teachers and between students and teachers.

Teachers' confidence in assessment has been explored by Goos and Hughes (2010), who conducted an on-line survey among more than 300 coordinators in an Australian university on the confidence level of assessment practice. They found that the co-ordinators were confident in making judgements of assessment but they felt less confident about the external requirements of their assessment performance.

Further evidence of markers' emotional reaction to the assessed work was found in studies on markers' thinking in the reading process of assessment by Crisp (2008). Crisp collected audio-recorded think–aloud data from six experienced examiners in geography who had practised marking and think-aloud in advance and who then marked four to six written scripts from two examinations. She found that the markers "sometimes showed like, dislike, amusement, frustration or other personal response to students' work" (p. 255). The markers assessed mostly according to criteria but they also demonstrated their reactions to the language of the written work, their assumptions about the students, and emotional reactions to the work, although these factors were not found to influence grading. However, Crisp only identified some evidence of markers' emotional reactions in the reading process of assessment. She did not explore the social aspects of emotion, or the interactions between emotion, cognition, and action in assessment.

To sum up, emotion and cognition interact with each other. Emotions such as empathy, trust, and confidence influence teacher cognition and their assessment practice. The results of assessment cause students' emotional response and affect their engagement with feedback on their work. It is suggested that feedback on positive aspects of written scripts may be able to reduce the possible negative emotional response of students. The roles emotion play in tutor cognition and in assessment feedback will be analysed in detail in this study.

2.3.3 Summary

In sum, three fundamental issues of teacher cognition studies are of current concern in the literature. First, there is a need for theoretical development that can reveal the interaction between cognition, emotion, and context. Secondly, there is a need for methodological development that can explore the complexity of teacher cognition. Thirdly, there is a need to explain the relation between emotion and cognition. In addition, there are gaps identified by the review of teacher cognition studies: a lack of research on novice teachers and the evolution of their thoughts; a need for studies on teacher cognition to incorporate literacy skills into specific content areas; a need to explain how university teachers develop theories of teaching; and a need to explain the role context plays in language teacher cognition and emotion. The review of assessment feedback studies that relate to teacher cognition and practices has also demonstrated the need for theoretical and methodological development. The main issues reviewed in this chapter, especially in regard to context, emotion, and methods in the study of teacher cognition, will be discussed from the perspective of a holistic approach to socio-cultural models of cognition.

2.4 Socio-cultural models of cognition

Conventional teacher cognition studies focus on the individual level. However, according to Vygotsky (1978), individual cognition has social determinants. This section will review socio-cultural theories that can be synthesized into a socio-cultural framework to address the existing issues in teacher cognition studies in general and tutor cognition of assessment feedback in particular. These theories include socio-cultural constructs derived from the seminal work of Vygotsky

(1978), theories of distributed cognition (e.g. Salomon, 1993), situated cognition (Lave & Wenger, 1991), and activity theory (Engeström, 1987).

2.4.1 Vygotsky's cultural-historical approach to cognition

A starting point to understanding Vygotsky's (1978) theory of cognition is the zone of proximal development (ZPD), a metaphor of the cognitive development mechanism. It refers to:

> …distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers. (p. 86)

The concept of ZPD is based on Vygotsky's central theory that human cognition is a collective and shared activity; therefore, cognition first happens between individuals on the social plane and then within individuals' minds. At a social level, individual learners, whether children or adult, who are in a ZPD need to be guided or regulated by more experienced people via physical tools and symbolic artefacts such as language (Lantolf & Appel, 1994). These cultural tools "mediate the relationships between people, between people and the physical world, and between people and their inner mental worlds" (Lantolf, 2006, p. 69).

Furthermore, Vygotsky (1981) argues that individuals internalize reality in their minds by inner speech to reach a new stage of cognition which, in turn, is externalized to regulate their behaviour and finally achieve the transition from other-regulation to self-regulation.

According to Vygotsky (1987), a ZPD cannot be diagnosed by formal testing or measurement; instead, it requires collection and analysis of multiple sources. This point of Vygotsky has been applied by Poehner and Lantolf (2010) to research on assessment in second language education. They advocate the principle of dynamic assessment. This principle refers to "the dialectical unity of instruction and assessment" (p. 312). It is integration of ongoing assessment and guidance on the basis of interaction and negotiation between learners and teachers (Poehner & Lantolf, 2010).

In addition, Vygotsky (1986) argues that there is dialectical relationship between cognition and emotion:

> Thought is not begotten by thought; it is engendered by motivation, i.e. our desires, our interests and emotions. Behind every thought there is an affective-volitional tendency which holds the answer to the last 'why' in the analysis of thinking. A true and full understanding of another's thought is possible only when we understand its volitional basis. (p. 252)

However, Vygotsky (1978) did not provide detailed analysis on how emotion and cognition interact with each other. Emotion has been largely neglected in studies of cognition using Vygotsky's psychological theory. This absence has been noticed only by a few researchers like Wells (1999), who argues that "Learning in the zpd involves all aspects of the learner-- acting, thinking and *feeling*" (p. 331 - emphasis added).

Vygotsky's (1978) socio-historical theories of cognition, especially the concepts of ZPD, mediation, and regulation, have rarely been applied to studies of teacher or tutor cognition at tertiary level, which is the research space occupied by the present study. Vygotsky's socio-historical theories laid the foundation for three concepts: distributed cognition, situated cognition, and activity theory. These three concepts have their specific definition of context and unit of analysis (Nardi, 1996), and thus are relevant to the present study. These three concepts will be reviewed in the following sections.

2.4.2 Distributed cognition

Vygotsyky's (1978) cultural-historical theory of cognition expands the unit of analysis of individual cognition to social dimensions. This social dimension of cognition is applied by Hutchins (1995) to cognitive science and interpreted into the concept of distributed cognition, describing how cognition is distributed among a group of individuals in a work setting to carry out tasks. The unit of analysis in distributed cognition is "a collection of individuals and artefacts and their relations to each other in a particular work practice" (Rogers & Ellis, 1994). Distributed cognition emphasizes the roles structure or system play in the alignment among a group of individuals in a shared process to achieve a collective goal (Nardi, 1996).

The concept of distributed cognition is also applied to the educational context by the advocates of activity theory: according to Cole and Engeström (1993), a cultural-historical approach to distributed cognition analyses the interactions "between an individual, a mediating artifact, and the environment" (p. 17) within an activity system across time. On the other hand, Salomon (1993) argues that the study of distributed cognition should not ignore the individual dimension. The concept of distributed cognition is applied to the present study in the aspect that it analyses interactions between individuals as a group and the tools they use in the working settings.

2.4.3 Situated cognition

Situated cognition, like distributed cognition, is another approach to the study of the social dimensions of cognition. It emphasizes that knowledge comes from action in certain contexts (Brown, Collins, & Duguid, 1989; Lave & Wenger, 1991; Greeno, 1998). The unit of analysis of situated cognition is the relation between the individual and the settings where they act (Nardi, 1996).

Lave and Wenger (1991) develop the situated approaches of cognition by introducing the notion of apprenticeship in a community of practice. According to Lave and Wenger, learners' cognitive development is always situated in a context where learners, as apprentices, participate in a community of practice and regard a community of practice as "an intrinsic condition for the existence of knowledge" (p. 98). They explain that:

> In using the term community, we do not imply some primordial culture-sharing entity. We assume that members have different interests, make diverse contributions to activity, and hold varied view points. In our view, participation at multiple levels is entailed in membership in a community of practice… It does imply participation in an activity system about which participants share understandings concerning what they are doing and what this means in their lives and for their communities. (pp. 97-98)

Lave and Wenger (1991) explain that the nature of the participation of newcomers in a community of practice is legitimate peripheral participation, which means the novices are allowed to undertake tasks that are less difficult and less risky: these tasks are increasingly complex as the newcomers demonstrate their enhanced

knowledge and/or skills. Furthermore, they also argue that the legitimate peripheral participation causes contradictions in a community of practice:

> Granting legitimate participation to newcomers with their own viewpoints introduces into any community of practice all the tensions of the continuity-displacement contradiction. These may be muted, though not extinguished, by the differences of power between old-timers and newcomers. (p. 116)

This perspective of contradiction in the community of practice is closely related to that of activity theory, which will be explained in the following section.

2.4.4 Activity theory

The third social cognitive approach to be considered is activity theory, which has developed from Vygotsky's (1978) socio-cultural theory. It was firstly proposed by Leont'ev (1978) and then developed by Engeström (1987). Cole and Engeström (1993) use activity theory to explain how cognition is distributed among people by commonly shared mediators in an activity system. Activity theory provides a tool and philosophical model to analyse the interactions not only within one activity system but also between activity systems (Barnard, 2010). Therefore it provides the potential to analyse the interaction between individual cognition and extended systems of distributed cognition. It is this potential that makes activity theory an appropriate framework to illuminate the findings of this study.

This section will briefly review Leont'ev's model of activity theory and then focus on the expanded model by Engeström. It will finally review the limited number of studies on assessment activity in universities that have applied activity theory.

2.4.4.1 Leont'ev's model

The original model of activity theory was developed by Leont'ev in the early twentieth century but was published much later.
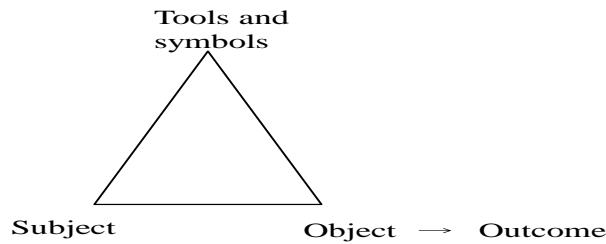
Figure 2.1: Original model of activity theory (Leont'ev, 1978)

Leont'ev (1978) explains his model using the example of a group of primitive people carrying out the activity of hunting. The subject in this model is a group of hunters. The object is the animal. With the help of both physical and cultural tools (primarily hunting equipment and language), the subject makes the transformation of the object. The outcome is the result of the transformation-the animal is killed. Leont'ev's model makes a clear distinction between activity, action, and operation, which form a three-level model of activity (see Figure 2.2).
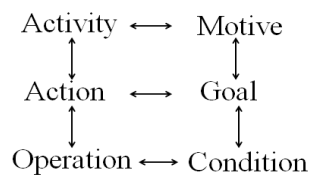
Figure 2.2: Leont'ev's (1978) hierarchical structure of activity system

Engeström and Miettenin (1999) provide a detailed explanation of this model:

> The uppermost level of collective activity is driven by an object-related motive; the middle level of individual or group action is driven by a goal; and the bottom level of automatic operations is driven by the conditions and tools of action at hand. (p. 4)

For example, an activity of assessing undergraduates' writing is carried out by teachers by undertaking a series of actions and operations to achieve the object – the provision of assessment feedback. The lowest level within this framework are operations such as underlining parts of the assignment by using the tools at hand, such as pens or pencils, or grading according to marking schedules. These operations enable the teacher to move to the higher level of actions, such as decision making or formulating ways of directing students' attention to errors and/or making improvements. Collectively, these operations and actions comprise

37

the overall activity. It is worth noting that, in most cases, operations need to be deliberately learnt as actions before they can be applied automatically; when operations are learnt as actions, there need to be subordinate operations and conditions to allow this to happen: for example, it is necessary to learn how to use the marking schedules.

2.4.4.2 Expanded model by Engeström

Engeström (1987, 1999a) argues that the original model of activity theory is not able to reflect the complex nature of human activity. Therefore, Engeström (1987) has expanded the original model into the following model of an activity system (Figure 2.3).
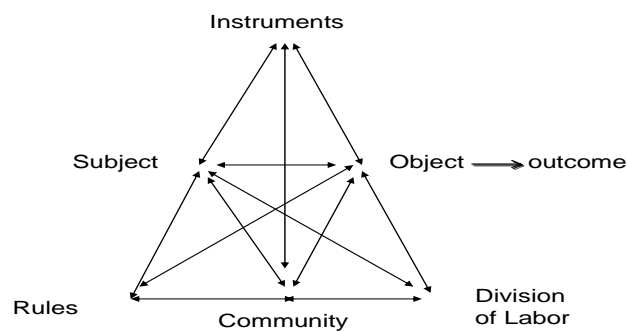
Figure 2.3: Engeström's (1987) model of an activity system (p. 78)

In the expanded model, Engeström (1987) includes more social components than Leont'ev's (1978) model, which form a second layer in the activity system. According to Engeström (1996), human activity is mediated by rules which include "regulations, norms and conventions that constrain actions and interactions" (p. 67). The work involved in the activity is distributed horizontally among participants, while status and power are divided vertically within the community – for example between tutors, lecturers and professors.

Engeström (2001) further summarizes activity theory into five principles: an activity as *the prime unit of analysis*, the *multi-voicedness*, *historicity*, *contradiction*, and *expansive transformations* of an activity system (pp. 136-137). The basic unit of analysis is an activity not only in terms of its inner interaction between subjects but also its interrelationship with other relevant activities. For example, an assessment activity in a university is related to other academic

activities such as teaching and research, and also to non-academic activities like resource management and policy-making. An activity system is multi-voiced because all the subjects carrying out the specific activity have different positions and carry different histories, as do the subjects in all other activities. A key point about Engeström's view of activity theory is that systems operate in time as well as space, and are therefore subject to change within an historical trajectory. Thus, the study of an activity requires analysis of the changes of an activity over time and the inevitable contradictions that arise from such change. Engeström (1987) has indicated four levels of contradictions. The primary level is within each component; for example, teachers who assess students' written work may want to facilitate students' cognitive improvement via formative feedback; however, they also have to fulfil the summative role of assessment. The secondary level is between components; for example, contradictions may occur between teachers who have different beliefs but use the same tool to carry out the activity of assessment. The tertiary level is between the existing model of an activity and that of a more advanced activity such as when a more effective model of assessment is developed. The quaternary level is between an activity and other activities within the same system; for example, the application of a new assessment model of activity may conflict with existing regulations relating to how assessment should be carried out.

According to Engeström (1987, 1999b, 2001), contradiction is the engine that motivates the development of both individual and collective cognition in learning activities because when these contradictions are recognized to alter the balance of the activity, the members of the community of practice need to make positive adjustments to their cognition. Such enhanced cognition begins to transform the old activity system to a new one, which Engeström (1987, 1999b, 2001) refers to as the cycle of 'expansive learning'.
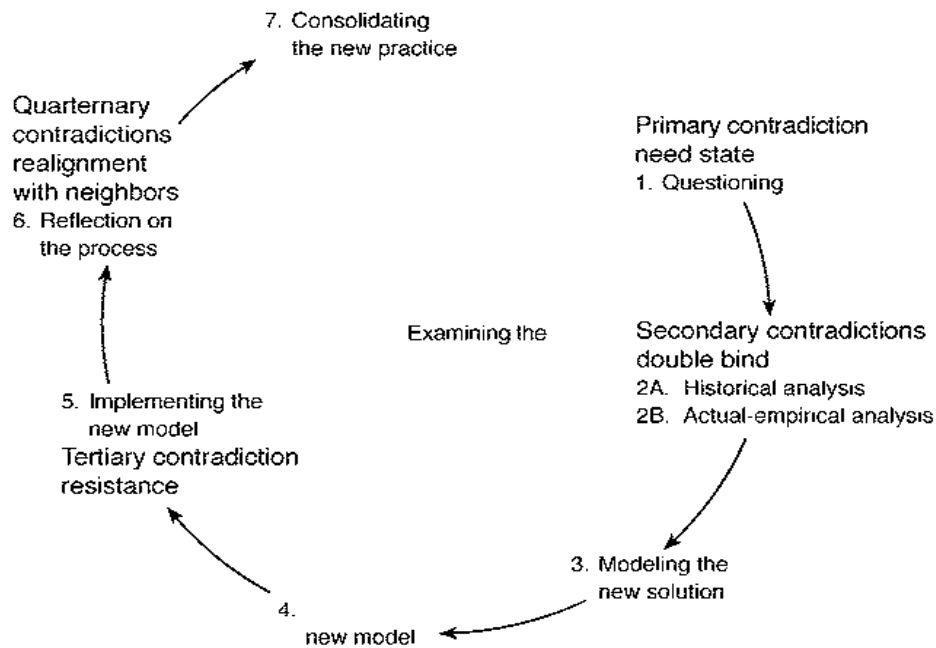
Figure 2.4: Engeström's (2001) model of strategic learning actions and corresponding contradictions in the cycle of expansive learning (p. 152)

As depicted in Figure 2.4, the strategic actions of expansive learning begin with questioning current practice by subjects and their need to make improvements. The second step is analysing the problems and contradictions within the current activity system. The consciousness of the problems in the current system as a result of analysis is first stage of cognitive transformation of subjects. Then subjects take actions to model new instruments, the process of which transforms the current activity into a new activity. However, the application of the new model requires efforts to reduce the tertiary level of contradiction between the new and the old activity. The transformation of the old activity also leads to quaternary level contradictions between the new activity system and its neighbouring activities. To reduce the quaternary level contradiction, collaboration between related activities is needed.

However, there are two issues of activity theory that need to be addressed. One issue is identified by Thompson (2004) who claims that Engeström's (1987) expanded model focuses on the interactions between organizational communities, which is a shift from Vygotsky's (1987) original focus on the role of social activity and interaction on individual cognitive development. The second issue of Thompson's (2004) argument is that there is a need for activity theory to explain

40

individual cognitive development in social activity, which point will be addressed in the present study.

Another issue to be addressed in activity theory is the role emotion plays in activity. Leont'ev (1978) mentioned emotion in relation to how well the motives correspond to it. However, with the exception of the study by Roth (2007), the emotional aspect of social activity has not been explored in depth in studies that apply activity theory. Roth, based on his longitudinal ethnographic study on workers in a hatchery, concludes that emotion influences both body states and decision making. Roth (2007) also claims that motivation and identity are mediated by emotion. The need for exploring the role of emotion in activity theory has been noticed by Daniels and Warmington (2007) who state that activity theory can be developed by the understanding of three issues: contradiction, the identity of the subjects in the activity system, and emotion. Engeström (2009) has also pointed out that one direction of the development of activity theory is "Moving down and inward, it tackles issues of subjectivity, experiencing, personal sense, emotion, embodiment, identity, and moral commitment" (p. 308).

The application of activity theory to the current study can contribute to the understanding of the cognitive development of tutors as subjects in the assessment activity as well as their assessment activity regarding the object of regulating students' cognitive development. The present study will contribute to activity theory regarding the connection between thinking and doing, and the relation between individual and collective levels of activity (Davydov, 1999). Most importantly, the present study will explain the interaction between emotion, cognition, and actions (ECA) in the activity of assessment, and thus contribute to the development of activity theory.

2.4.4.3 Activity theory approach to studies of assessing writing

Activity theory has been used by a very small number of studies on writing across the curriculum (Burton, 2010; Russell, 1997) and formative assessment (Crossouard, 2009; Crossouard & Pryor, 2008). Crossouard and Pryor are perhaps the first to apply activity theory to feedback and assessment. They studied theory and practice of formative assessment in a doctorate programme in UK by multi-methods of data collection among 11 doctorate candidates. They regarded

formative assessment as an activity system which included both institutional, disciplinary and other relevant communities with both students and teachers as subjects. The problem with this study is that the system they used combined too many sub-activity systems to be able to explain the relations and contradictions in the local community of practice. Cross (2010) proposed a socio-cultural approach to language teacher cognition study by applying Vygotsky's (1981) genetic analysis and Engeström's (1987) model of activity theory to the analysis of the relationship between cognition, practice, and context. He exemplified this approach by analysing survey data of language policy and the audio-recorded stimulated recall and interview data of a non-native teacher of Japanese. The core of this approach is the three-level analysis: cultural-historic domain of the policy context, ontogenetic domain of teachers' cognitive background, and micro-genetic domain of the teacher's beliefs and actions in relation to the key elements in the activity system. According to Cross (2010), the particular value of this approach was that it offered a framework to explain the role of context in teacher cognition and the contradiction in the activity system. Cross concluded that this socio-cultural approach provided a holistic frame for teacher cognition studies in that it synthesized thinking, doing, and context. The value of Cross's (2010) socio-cultural model lies in its integration of macro-level context (education policy) and micro-level context (the immediate activity system) of teacher cognition. However, there is a need to explain how each element in the activity system interacts with individual *thinking* in the process of *doing*, which Vygotsky (1978) termed 'cognition in flight'.

2.4.5 Summary

This section has reviewed Vygotsky's (1978) social-historical theories of cognition, distributed cognition, situated cognition, and activity theory. These models provide a holistic perspective on cognition studies which will be applied to the present study.

2.5 Summary of the chapter

Through the preceding review of theories and studies of teacher cognition and feedback, and socio-cultural theories, the following gaps have been identified in the current literature.

Firstly, studies on assessment feedback have been carried out in various sub-branches of studies in education, linguistics, and cognition; however, there has been no research that connects the branches of study together to provide a holistic understanding of subject teacher cognition of assessment feedback. Secondly, few studies have been carried out on the cognition of subject teachers, especially subject tutor cognition of giving assessment feedback on undergraduate written work. Thirdly, research on teacher cognition is mainly carried out on teachers' beliefs at the individual level without in-depth study on the sources of those beliefs. There is little research on how individual university teacher cognition is distributed or interacts within its community of practice. Fourthly, there is little evidence on how teachers' beliefs and practices interact with contextual factors, nor is there a clear definition of context in university teacher cognition studies. There is a research space for how context can be defined and used to provide a meaningful explanation of teacher cognition of giving assessment feedback. Fifthly, there is little research that applies a holistic Vygotskian perspective to the exploration of teacher/tutor cognition of giving assessment feedback. There is evidence that emotion influences cognition, but how emotion and action interact with teacher cognition is underexplored. Finally, current research methods on teacher cognition of giving assessment feedback are limited. For example, few studies have incorporated data collected by think-aloud methods during the natural practice of giving feedback into teacher cognition studies.

To sum up, thus far it has proved impossible to locate published research on subject tutors' beliefs about and practices of giving written feedback on students' written assignments in New Zealand universities. Those studies which have been published elsewhere are limited in their research focus and do not include the convergence and divergence of subject tutors' beliefs and practices in the actual activity of assessment, both as individuals and as members of specific communities of practice. Neither have they applied a holistic socio-cultural perspective, such as has been reviewed in this chapter, to address the contextual issues in the study and provide an interpretation of teacher cognition of assessment activity. Current research methods used in teacher cognition studies mainly rely on document analysis, survey, and interview, and these methods are inadequate to explore the dynamic relationship between individuals and their

community of practice. Moreover, the cognitive theoretical framework for current research on teachers' beliefs and practices is incomplete. Finally, there is a need to explain the function of emotion in teacher cognition and practice. Therefore, there is a need for this present study to investigate the questions of what tutors believe and do in providing assessment feedback, what factors influence their beliefs and practices, and how these issues can better be explored.

By applying a holistic socio-cultural approach to the study of tutor cognition of assessment, this research aims to contribute to both theoretical and methodological construction of teacher cognition studies with practical implications for subject tutors' practice of assessment, and professional and cognitive development of tutors. Tutors' cognitive development in the activity of assessment will be explored both at individual and collective levels in relation to emotion and action.

The next chapter, Methodology, will explain the data collection and analysis methods used to achieve the aims of the study.

CHAPTER THREE: METHODOLOGY

3.0 Introduction

This study is an inquiry into teacher cognition and its relation to the practical activity of assessing discipline-specific writing. It is an investigation, guided by activity theory, about teachers' beliefs and practices in their natural working contexts. It takes interpretive naturalistic inquiry as its paradigm and primarily adopts a qualitative approach to data collection and a grounded approach to data analysis. The study of teacher cognition and activity is a complex issue, which by its nature requires an open-minded philosophical foundation and a combination of different methods of data collection, analysis, and interpretation to address the following research questions.

- --What do subject tutors in the specific context believe about giving written feedback on students' assignments?

- --What are their actual practices when giving feedback?

- --To what extent do their actual practices converge with or diverge from their beliefs, individually and across departments/disciplines?

- --What are the socio-cultural factors that influence tutors' beliefs and practices?

- --How do the findings to these questions add to academic understanding of what constitutes teachers' beliefs, and the possible tension that arises in putting these beliefs into practice?

In order to best address the research questions, the activity of providing assessment feedback is taken as the unit of analysis. In chronological order, the following means of data collection were used: survey, interview, think-aloud, stimulated recall, and focus group discussion. The combination of these methods of data collection can not only reveal tutors' beliefs individually and as a group, but can also reveal what tutors do and think in the process of their practice and the factors that interact with their beliefs and practices. The analysis of data is informed by grounded theory, which best fits in the interpretive and naturalistic framework of this qualitative case study.

This chapter will introduce the methodology of the study. Section 3.1 will explain that this investigation is a case study under the socio-constructivist interpretive paradigm and uses a qualitative approach to data collection and analysis. The next section will analyse why and how a multi-method approach to data collection is needed for this research. Section 3.3 will discuss the ethical considerations and procedures of the study, and the position of the researcher. Section 3.4 will explain the grounded theory approach adopted for data analysis. This will be followed by an explanation of quality considerations. Sections 3.6 and 3.7 will describe the actual procedures of data collection and analysis, and the chapter will conclude with a summary of the key points raised and a brief preview of the following chapter.

3.1 Theoretical background of the methodology

This section introduces the theoretical background of the methodology. It includes three aspects: the socio-constructive interpretive paradigm, case study, and a qualitative approach.

3.1.1 Socio-constructivist interpretive paradigm

This research project is an inquiry into teacher cognition and its relation to practical activities associated with teaching, specifically the provision of feedback on written academic assignments. The study adheres to the socio-constructivist interpretive paradigm for the reasons discussed in the following section.

Firstly, it is interpretive because it is a small scale research project by the researcher who tries to gain insider knowledge of the subjective world of participants in their natural context. It considers both visible practice and invisible cognitive processes. It fits the interpretive paradigm that focuses on the understanding of "the subjective world of human experience" (Cohen, Manion & Morrison, 2007, p. 21). In other words, an interpretive paradigm takes reality as a subjective complexity, which is dynamically constructed and interpreted by human beings in their daily activities. As teachers' practices or actions are guided by their beliefs, their actions are intentional and meaningful. This is a point emphasized in the interpretive paradigm:

> Actions are meaningful to us only in so far as we are able to ascertain the intentions of actors to share their experiences. (Cohen et al., 2007, p. 21)

The primary aim of the interpretive researcher is to gain understanding and explore categories and patterns which emerge from analysis of the data and thereby derive grounded theories "from particular situations" (Cohen et al., 2000, p. 23). Such an exploration is different from normative research which seeks a universal theory, or critical research which seeks to enhance the emancipation and critical consciousness of participants.

Secondly, it is supported by a Vygotskian socio-cultural perspective and takes its epistemological perspective by regarding the meaning of reality as co-constructed by individuals in social contexts (Vygotsky, 1978). In this research context, the beliefs and practices of giving written feedback are distributed within and among teachers and tutors at the faculty level but the study takes into account the wider socio-cultural context within which these teachers/tutors operate.

However, the key weak point of the interpretive paradigm is the "relative neglect of the power of external-structural-forces to shape behaviour and event" (Cohen et al., 2000, p. 27). Therefore, activity theory, a philosophical framework for research, is used to compensate for this limitation. According to Engeström (2001), the primary unit of analysis of a human activity is "a collective, artifact-mediated and object-oriented activity system, seen in its network relations to other activity systems" (p. 136). In the case of the current study, the tutors' beliefs and practices are studied through the lens of activity theory so that the influence of other components in the activity system on tutors' beliefs and practices can be revealed. In this way, a more thorough understanding of the participants' activity and cognition can be gained.

3.1.2 Case study

Hood (2009) defines a case as "a bounded system comprised of an individual, institution, or entity and the site and context in which social action takes place, the boundaries of which may not be clear and are determined by the scope of the researcher's interests" (p. 69). According to Duff (2008), case studies have the following attributes: "boundedness or singularity, in-depth study, multiple perspectives or triangulation, particularity, contextualization, and interpretation" (p. 23). It is regarded as the method "most suited" (Cohen et al., 2000, p. 181) to

interpretive paradigm and a method "particularly appropriate for individual researchers because it gives an opportunity for one aspect of a problem to be studied in some depth within a limited time scale" (Bell, 1999, p. 10). It is especially suitable to explore the "how and why of a complex situation" (Yin, 1994, p. 16) or "complex social units consisting of multiple variables of potential importance in understanding the phenomenon" (Merriam, 1998, p. 41). Case study also has the advantages of allowing multi-methods of data collection (Creswell, 2007; Dörnyei, 2007; Duff, 2008; Hood, 2009; Merriam, 1998; Stake, 2005; Yin, 2003).

The first issue a case study needs to address is how to define specific boundaries of the case (Creswell, 2007; Duff, 2008; Yin, 1994, 2006). Traditionally, a case is regarded as "an integrated system" (Stake, 1995, p. 2) and a bounded unit of analysis (Creswell, 2007; Dörnyei, 2007; Duff, 2008; Hood, 2009; Merriam, 1998; Stake, 1995; Yin, 2003), which results in the following three issues: firstly, how to define the boundary of a case; secondly, how to analyse this unit as a system; and thirdly, whether the system interacts with its context and other systems. Stake (1995, 2005) did not define the term *system* or *boundary*. Merriam (1998), however, bounded a case by "a limit to the number of people involved who could be interviewed or a finite amount of time for observations" (pp. 27-28). Hood (2009) made a clearer boundary by stating that a "bounded system is composed of an individual (or institution) and a site, including the contextual features that inform the relationship between the two" (pp. 68-69).

In the current study, the case is the activity of providing assessment feedback by individual tutors and as a group. As explained in Section 2.4.4, the case is analysed through the lens of activity theory.
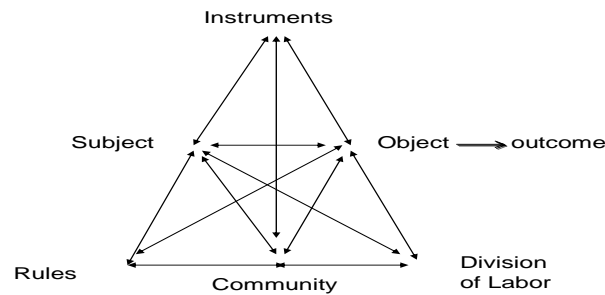
Figure 2.3: Engeström's model of an activity system (1987, p. 78)

The model is presented here again to explain the framework for the contextual analysis of the activity. By analysing tutor cognition in relation to the contextual elements, relationship between individual and distributed cognition of the activity can be discovered. The concept of the three levels of activity (i.e. activity, action and operation as has been explained in Section 2.4.4.1) contributes to a comprehensive and systematic in-depth collection and analysis of data. Moreover, the boundary limits are not permanent fixtures but are only identified for the sake of study. Nothing in the activity system remains unchanged because any activity system is constantly subject to change (Engeström, 1999b) resulting from internal or external pressures. Therefore, change is significant in this study. This perspective is different from conventional studies that either regard change as one of the disadvantages of case study or leave the issue of change unaddressed (Duff, 2008). The following sections will address the issues of data collection and analysis by overviewing the qualitative approach, multiple methods of data collection, and the grounded theory approach to data analysis.

3.1.3 Qualitative approach

A qualitative approach is preferred to a quantitative approach when the purpose of research is to

> provide perspective rather than truth, empirical assessment of
> local decision makers' theories of action rather than generation
> and verification of universal theories, and context-bound
> explorations rather than generalizations. (Patton, 1990, p. 491)

This research project adopts a qualitative approach because it aims to explore beliefs and actions in the natural context rather than argue for universal truth or theories. It fits the four characteristics of qualitative research summarized by Merriam (2002):

> To understand the meaning people have constructed about their world and their experiences;
> The researcher is the primary instrument for data collection and data analysis;
> The process is inductive; and
> The product of a qualitative inquiry is richly descriptive. (pp. 4-5)

Qualitative research aims understand peoples' beliefs and practices, which is the aim of the present study. The people to be studied are tutors working in various departments within a faculty of a university in New Zealand. The tutors belong to different local communities of practice within a large activity system of assessment in the university, yet they share many of the same students. Their common function is to help students improve their writing by giving feedback. In doing so, they have to abide by some common academic, pedagogical, and institutional rules and share some academic and pedagogical knowledge of good practice of giving feedback. The exploration of the beliefs and practices of the tutors, both as individuals and as a community in the activity system, requires qualitative approach of data collection, which will be explained in the next section.

In addition, the project also takes advantage of key features of qualitative research summarized by Johnson and Onwuegbuzie (2004) and Strauss and Corbin (1990), such as identifying key contextual factors and addressing complexity, providing detailed insider description of dynamic process, gathering rich and in-depth data , exploring possible causes of the phenomena, and generating grounded theory.

3.2 Multi-method approach of data collection

Kagan (1990) suggested that a multi-method approach of data collection should be used in teacher cognition studies as a methodological triangulation to achieve internal validity.

This qualitative case study takes a multi-method approach in its data collection, which comprises survey (using both closed and open-ended questions), semi-

structured individual interview, think-aloud, stimulated recall, and focus group discussion; these are among the procedures currently adopted to explore language teacher cognition (Borg, 2006). The following section will explain the methodological considerations of each data collection method.

3.2.1.1 Survey

Survey, the most commonly used method in teacher cognition research (Borg, 2006), is basically a method designed to collect quantitative data for statistical analysis (Fowler, 2009). This method has its advantages, being quick and economical, but data collected by this way "are obviously limited in their ability to capture the complex nature of teachers' mental lives" (Borg, 2006, p. 174). Therefore, it is often used together with additional qualitative methods.

The e-questionnaire (Appendix C) was used at the beginning of this study for four purposes: to collect some bio-data background about the tutors across the faculty; to obtain a 'snapshot' (Nunan, 1992) of tutors' attitudes about giving feedback on written assignments at department and faculty level; to provide a baseline for more in-depth research; and to invite participation in a later phase of data collection. The e-questionnaire followed the general format suggested by Dörnyei (2007). The content comprised "Factual questions, behavioural questions, and attitudes questions" (Dörnyei, 2007, p. 102). The closed items had three forms: Likert scales, true or false items, and multiple-choice questions. It also included opportunities for open-ended comments.

3.2.1.2 Semi-structured interview

The semi-structured interview (Patton, 1990) is a method frequently used in studies to flexibly elicit and probe teacher cognition of giving feedback (e.g. Borg, 1998; Lea & Street, 2000), either as the only research tool or together with other instruments (Bogdan & Biklen, 1982).

Merriam (1998) notes two key qualities of interviewing: questioning and interaction. Multiple, leading, and polar questions should be avoided (Merriam), and good questioning and probing skills are based on practical experience. Therefore, before collecting interview data, researchers should practice being both interviewer and interviewee, and continually reflect on their interviewing manner

and skills to "get to know themselves" (Johnson, 2002) as an interviewer; this will be discussed in Section 3.6.1. The quality of interaction can be facilitated by a relaxed setting, friendly manner of interviewing, and establishing the "insider-outsider status" (Merriam, 1998, p. 86) between the researcher and the participants.

Semi-structured interviews were used in this study for five purposes. The first was to generate data about individual tutors' beliefs and their reported practices in giving assessment feedback. The second was to explore the sources of their beliefs. The third was to find common beliefs about giving feedback among tutors. The fourth was for the interview data to be used in triangulation with think-aloud and stimulated recall data to identify possible convergences and divergences between beliefs and practices of individual teachers/tutors. The fifth was that the data could be used to compare the convergence and divergence between the beliefs of individuals and those of their community of practice.

The focus points of the semi-structured interview (Appendix E) were designed on the basis of the research questions and informed by the survey data. Using focus points in a semi-structured interview, rather than a series of interrogatives, allowed the flexibility to explore relevant data emerging during the oral interaction. My role in the interview was that of a facilitator (Cohen et al., 2000) to motivate participants to address the focus issues as fully as possible (Brenner, 2006). Interviews were audio-recorded and transcribed, and summaries were sent to the interviewees for respondent validation.

Both surveys and interviews are conventional procedures widely discussed in textbooks of methodology or applied in empirical studies. On the other hand, think-aloud and stimulated recall sessions are much less established in teacher cognition studies, so more attention will be paid to these two methods in the following sections.

3.2.1.3 Think-aloud
Think-aloud is a method to collect cognitive data by asking participants to verbalise their cognitive process during the process of performing tasks (Ericsson & Simon, 1984; Gass & Mackey, 2000; Mckay, 2009).

According to Ericsson and Simon (1984), collecting think-aloud data involves several basic points. First, before conducting think-aloud, researchers should select one or more tasks, train participants, and provide instructions for the session. Second, during the process of think-aloud, researchers should not communicate with the participants except to remind participants to keep on talking. Finally, think-aloud verbalization has three levels: level one is the direct verbalization of thought; level two is verbalization and description of thought; and level three is "an explanation of thought" (p. 79). Ericsson and Simon believed that the first two levels of data should be explored to find the concurrent thinking process while level three data is invalid for the study of thinking process because it relates to previous thought.

Ericsson and Simon's (1984) approach of conducting think-aloud has been widely used in applied linguistics in studies of the language processing of students (Mckay, 2009) and has been used in a limited number of studies on language teacher cognition (Borg, 2006). However, this approach is criticized for it "is often used in conditions especially created for the purpose of research" (Borg, 2006, p. 224). This was indeed the case in the studies by Cohen and Cavalcanti (1990) and Barkaoui (2007). The main issue in contention is that Ericsson and Simon's (1984) approach is rooted in information processing theory that focuses on how language can merely reflect the thinking process. This is in contrast to Vygotsky's (1962) theory that thinking is mediated by language. Based on Vygotsky's (1962) social historical psychology, Smagorinsky (1998) argues that think-aloud should always be socially situated and audience-oriented. In agreement with Smagorinsky (1998), Swain (2006) argues that verbalization is not a mere reflection of thought but a process of shaping and sharpening thought. In addition to theoretical arguments about think-aloud, there have been procedural changes to Ericsson and Simon's (1984) approach regarding the researcher's role in obtaining think-aloud data. For example, Cooksey et al. (2007) used "think aloud dialogue" (p. 409) in their study of writing assessment. One interesting point of their study is that it promoted an interactive think-aloud between researchers and participants rather than a pure monologue. This is in agreement with Boren and Ramey (2000), who claim that little data could be obtained if researchers remain entirely in the background. Similarly, Charter (2003) insists

that think-aloud should be conducted in a natural way. She used audio recording and sat by her student participants, trying to "keep their think-aloud behaviour as natural as possible" (p. 74). However, the specific procedures of conducting qualitative think-aloud in natural settings have not been systematically developed perhaps due to "serious logistical difficulties" (Borg, 2006, p. 224) when used in natural settings. Such difficulties include practical issues such as access to participants, ethical issues, time negotiation, and so on.

Therefore, there is a need for qualitative researchers to address the procedural issues of conducting think-aloud in natural settings; such an attempt was made in this study. The use of think-aloud in this study is different from Ericsson and Simon's (1984) model in both theoretical and contextual aspects. Theoretically, the think-aloud approach in this study is based on Vygotsky's (1962, 1978) theories that language is a tool mediating thinking and that thinking happens first at the social, then the individual, plane. Therefore, the explanation of thinking, which is grouped by Ericsson and Simon (1984) as the third level data, is also regarded as valid in the present study. Moreover, the thinking process should reflect the interaction between the social and individual planes. This was regarded as valuable data in this study that could be obtained by think-aloud. Contextually, the think-aloud procedures adopted in the present study are different from Ericsson and Simon's (1984) model in four aspects. First, the purposes of conducting think-aloud after the interview were for me to learn how participants assess students' written work and give feedback in real context, and to explore factors that interact with their on-going cognition. Second, the task to be thought aloud was not selected by the researcher specifically for the purpose of gathering data from individuals. Instead, the task was tutors' routine work, which was not designed for the purpose of this present study but had actual social effects on the tutors, the students whose work was assessed, and the faculty. Third, tutors' normal work of assessment, generally speaking, involved three actions: reading and comprehending the assignment, evaluating the written work, and finally writing up the feedback (encoding what they thought about the assignment in written language). The participants in this study were expected to think-aloud the cognitive process while doing the above. The verbalization of the process was expected to involve the first two levels of Ericsson and Simon's (1984) model, but it may also involve an *a posterior* explanation of the decision participants make

54

because feedback itself was an evaluation of the written work and a justification of the grade, which fell into Ericsson and Simon's level three data. For this reason, some level three data were also valid in this study. Finally, Smagorinsky (1998) argues that think-aloud is never isolated from its social settings and its audiences. In the current study, the audiences were not only tutors themselves (in the sense that they were engaging in internal conversations) , but also included the students who were the target addressees of assessment feedback, the lecturers who might review tutors' work, and the researcher (myself) who would analyse the content of think-aloud. This was different from Ericsson and Simon's (1984) model, which ignored the existence of audiences.

Based on the purpose and context of think-aloud in this study, the procedure of conducting think-aloud was designed as follows. Firstly, instructions for think-aloud were provided to participants before the actual session (Appendix F). Negotiations were made between participants and the researcher on whether or not some form of training was needed. Secondly, participants were expected to verbalise their thought processes while marking assignments in the presence of the researcher. The researcher had the following roles in the think-aloud sessions: to motivate participants to think-aloud by quietly back channelling when necessary or simply by her social presence; to observe the process of assessment and keep field notes on decisions, especially the non-verbalised ones, that participants made in the process of assessment which were to be explored in the subsequent stimulated recall session. Finally, the setting of the think-aloud in this study was tutors' work setting rather than in laboratories. The time, type of assignments, and place of doing think-aloud were all decided by the natural working context of tutor participants. The think-aloud could only be conducted when participants had assignments to mark and were willing to think-aloud their cognitive process in the presence of the researcher (Li & Farrell, in press).

The advantage of using think-aloud in this present study is that it can better reveal tutors' thinking process and the factors that affect the thinking process while providing assessment feedback. The disadvantage is that think-aloud may increase the cognitive load of participants and distract their attention from their work (Li & Farrell, in press).

3.2.1.4 Stimulated recall

As think-aloud cannot in itself provide complete data (Ericsson & Simon, 1984), it is rarely used as the only method of data collection (Charter, 2003; Wilson, 1994). Therefore, "retrospective verbalization" of thought after task performance (Ericsson & Simon, 1980, p. 220) or stimulated recall (Gass & Mackey, 2000) was employed in the present study to compensate for the incompleteness of data**.**

Currently, the combination of think-aloud and stimulated recall is not common in teacher cognition research. However, one example is the study carried out by Scarino (2005), who spent two years collecting both introspective and retrospective data from three teachers of French in a secondary school in Australia. Introspective methods were used when teachers were assessing different types of written assignments of their students. Then a week later, the tape recording of the introspection session was used as stimulus in retrospective sessions. Unfortunately, no information was given by Scarino (2005) on the detailed procedures of conducting the two sessions.

A few examples of the combination of the two methods can be found in second language research. For example, Haastrup (1987) used combined think-aloud and retrospective interview methods for 32 pairs of student participants. To generate more data, students were asked to do think-aloud in pairs. The process was visually monitored by two researchers in another room who took notes on the process. Retrospective interview was used immediately after the think-aloud between each researcher and student. Haastrup concludes that combining introspective and retrospective methods compensates for the shortcomings of each. In addition, it is necessary for researchers to be present at the think-aloud session by listening to the verbalization and taking notes if they want to "use the retrospection as a complementary method" immediately afterwards (p. 211). This study reveals two key points that often present a dilemma for researchers who combine think-aloud with stimulated recall sessions: the temporal connection of the two sessions and the use of appropriate stimuli. According to Gass and Mackey (2000), stimulated recall sessions are most effectively conducted immediately after think–aloud sessions. However, Borg (2006) suggests that, because it takes time to select and transcribe even short elements of task performance, it is impossible to use transcripts to conduct stimulated recall

immediately afterwards. Borg (2006) also notes that the use of video may raise ethical issues and can be more intrusive if the task is carried out in the real context; also, it takes tutors extra time to listen to or watch recordings during the stimulated recall sessions.

Solutions to these issues were sought in the design of the stimulated recall in this study. The stimulated recall sessions were intended to compensate for non-verbalised decision-making processes by triangulating the think-aloud data with the participants' retrospective comments. This also gave participants an opportunity to further elaborate their beliefs about assessment and feedback. The instructions for stimulated recall (Appendix F) were e-mailed to participants together with those of the think-aloud sessions, so that participants could choose the appropriate time and place. Stimulated recall was carried out immediately after think-aloud sessions or, in one case, the following day. The topics for stimulated recall sessions were points selected by the researcher while observing the think-aloud sessions. The stimuli were the newly assessed written work, the criteria or marking schedules, and other tools the tutors used such as feedback sheets, and notebooks. All stimulated recall sessions were audio recorded.

3.2.1.5 Focus group discussions

A focus group is often regarded as "a discussion-based interview that produces a particular type of qualitative data generated via group interaction" (Lynne, 2006) – in other words, a group interview. However, Parker and Tritter (2006) argue that the contemporary use of focus group method is different from group interview in that:

> the researcher takes a peripheral, rather than a centre-stage role for the simple reason that it is the inter-relational dynamics of the participants that are important, not the relationship between researcher and researched. (p. 26)

The focus group method is used in order to gain "access to a sense of participant commonality" (Parker & Tritter, 2006, p. 24). It can be used either independently or together with other methods (Dörnyei, 2007; Li & Barnard, 2009; Morgan, 1997).

There are three key issues in using a focus group to collect data: the composition of group, the focus of the discussion, and the researcher's role in the focus-group interview. According to Stewart, Shamdasani, and Rook (2007), people form a group by "group cohesiveness" (p. 25) which is affected by their expectations of other group members, their "similarity of backgrounds and attitudes" (p. 25), and "the degree and nature of communication among group members" (p. 25). The interactions in the focus group are "akin to those that occur in everyday life but with greater focus" (Kamberrelis & Dimitriadis, 2005, p. 904). It is the interaction between group members that gives the advantage of this method (Dörnyei, 2007; Kamberrelis & Dimitriadis, 2005; Morgan, 1997; Stewart et al., 2007). The topics are usually questions provided by the researchers, either in the form of specific questions asked with time control or general questions that guide discussions with flexible amount of time (Morgan, 2002).

As a research method used in the final phase of data collection in this project, focus group discussion had the purposes to confirm and probe key points of data from individual interviews, to examine the notion of distributed cognition in communities, and to extend previous data by further exploration and co-construction of issues through group interaction.

The focus group discussion method was designed to include several criteria. First, tutors who had participated in some or all previous sessions of data collection were invited for focus group discussions, and seven actually took part. Second, the topics for focus group discussion were based on the emergent themes of the previous data. These topics of discussion were e-mailed to participants as part of the invitation letter (Appendix I). Hard copies of the topics were also provided to participants before the focus group discussion. Third, the researcher was not present in the focus group discussion. This is different from conventional role of researchers who are supposed to be the moderators of focus group discussion (Dörnyei, 2007; Kamberrelis & Dimitriadis, 2005; Morgan, 1997). The absence of the researcher in the focus group has two advantages: participants could exchange opinions on the given topics with their colleagues without being regulated and influenced by the researcher who was an outsider of the community, which further ensured the appropriateness of the composition of the focus group;

and participants had the flexibility in selecting topics and time control, which gave them the opportunity to elaborate the issues that most interested them.

The main advantage of using focus group discussion in the present study is that it reveals the common concerns among the members of the community of practice by the natural interaction between tutors. The focus group discussions at the end of data collection further triangulated the data already collected.

3.3 Ethical considerations

Research ethics influence the quality of research (Groundwater-Smith & Mockler, 2007). This section will introduce two aspects of ethical considerations: data collection and the position of the researcher.

3.3.1 Ethical considerations of data collection

The study was designed and carried out strictly according to the University of Waikato's Ethical Conduct in Human Research and Related Activities Regulations 2008. The general ethical considerations were given to voluntary participation, obtaining informed consent for the study including consideration of potential risks for participants and conflicts of interest and confidentiality, using pseudonyms of all participants, the departments, and the institution, and abiding by The Treaty of Waitangi (Appendix A).

Specific ethical consideration was also given to data collection within each method. For the e-survey, the tutors were approached via their email addresses, obtained from each department of the faculty, and invited to submit their anonymous responses via web-link to my e-mail address. The language and structure of the survey was designed to avoid taking too much of the participants' time, and the web-link ensured easy and quick access. In addition, the cover page was written to give clear guidance on how to complete the questionnaire; the cover page (Appendix B) also informed those who were willing to participate in follow-up sessions to contact the researcher separately via e-mail.

Ethical consideration of the subsequent sessions included the protection of participants' anonymity, negotiating time and place with participants before data

collection, using a digital audio recorder rather than video camera to avoid intrusion, minimizing interruption into participants' work due to data collection, and collecting data in a friendly atmosphere. Specific consideration was given to think-aloud and stimulated recall sessions by distribution of the letters of informed consent to departmental chairpersons and lecturers who oversaw tutors' work, providing them with information about the research, and seeking their approval for the use of their materials related to marking. (Appendix D)

The key ethical consideration of the focus group was the recruitment of participants. In order to have the informed consent of participants, the researcher must ensure that the composition of the focus group is appropriate for participants (Parker & Tritter, 2006). This issue was heuristically addressed in this study by recruiting participants from among those who had participated in previous stages. The participants of the focus group discussions were tutors who had participated in previous sessions. Therefore, they were likely to be interested in discussing assessment-related issues of their work with their colleagues.

3.3.2 Position of the researcher

Ethical consideration was also given to myself as the subjective agent of data collection, analysis, and interpretation. My previous identity was a university teacher before I came to New Zealand. My identity in the period of data collection was as a full time PhD candidate, and I was also working as a part-time tutor for international students at the faculty of my research context. Therefore, I was a legitimate, if peripheral, member of the community of practice of assessment of my research focus. The teaching, tutoring and learning experiences permitted me an emic understanding of the complex context of the research. However, my identity was that of a PhD candidate who had no power relationship with participants, and I was not a member of their actual speech community of practice. Therefore, I could retain a relatively impartial and etic standpoint in my data collection, analysis and interpretation. In addition, I had fresh sensitivity to factors that might be taken for granted by members in the centre of the focused communities in my research context because I was from another cultural background. Finally, my identity as a learner rather than an expert may have made participants more at ease in my presence thus helping me to obtain richer data.

However, I was also clear that I might have bias in my research. This bias may come from my cultural background, or from my experience. In order to overcome this limitation, I needed to be very cautious in the process of research; and an important way to avoid bias was to rigorously and systematically triangulate the data that were collected.

3.4 Grounded theory approach to data analysis

Grounded theory is an approach of collecting and analysing qualitative data, which was originally established by Glaser and Strauss (1967). Its main feature is to generate theory from data inductively by concurrent data collection and analysis, and constant comparison between data. After its first establishment, grounded theory fell into the contrasting Glaserian and the Straussian traditions (Charmaz, 2006; Creswell, 2007). The main argument between the two traditions is how to reconcile the contradiction between the theoretical background of research and the emergence of theories (Kelle, 2005). Strauss and Corbin (1990) suggest using axial coding, whereas Glaser (1978, 1992) insists on using the pure inductive approach and the "interpretive, contextual and emergent nature of theory development" (Burden & Roodt, 2007, p. 13).

Regarding the process of data analysis, Charmaz (2006) seems to agree that it is difficult for grounded theorists to completely avoid a deductive approach in data analysis. The general coding process she suggests includes initial open coding, focused coding, axial coding and theoretical sampling. This process can also be technically carried out by applying compute software such as NVivo8, which, according to Bazeley (2007), can greatly facilitate data organization and analysis. Bryant and Charmaz (2007) notice that there has been an increasing number of grounded theorists using computer software; however, they argue that researchers should take control of data rather than depending mainly on software.

In addition to the argument above on the process of data analysis, Braun and Clarke (2006) argue that there is a need for a practical approach of data analysis that stands on its own rather than being bounded by a theoretical position, such as grounded theory. Therefore, they have developed a qualitative data analysis method, thematic analysis, for "identifying, analysing and reporting patterns

(themes) within data" (p. 79). Thematic analysis follows the Strauss and Corbin (1990) tradition and acknowledges that themes of data can be found inductively but this process is also influenced in a deductive way by research questions. Braun and Clarke (2006) also suggested six steps of thematic analysis: "Familiarizing yourself with your data; generating initial codes; searching for themes; reviewing themes; defining and naming themes; producing the report (p. 87). These general steps were adopted in my process of data analysis with particular focus on the procedures of initial coding, categorizing, and generating themes.

In sum, grounded theory has historically developed and has been using as a main approach of qualitative data analysis. The grounded theory approach of making constant comparison and inductively generating theories has been established as general guidelines for qualitative data analysis; however, the detailed procedures of data analysis may vary especially regarding whether, and the extent to which computer software is applied. The analysis of data in the present study took grounded inductive approach with the assistance of NVivo 8which process is to be discussed in Section 3.7.

3.5 Quality consideration

According to Lincoln and Guba (1985), a qualitative study is trustworthy if it is credible, transferable, dependable and confirmable. These criteria have been widely quoted to justify the rigor of qualitative research.

Lincon and Guba (1985) argue that the most important criterion is credibility, which means that the study can present participants' world through their own eyes. Merriam (2002) listed five strategies to achieve credibility: member-checking, peer-reviewing, researchers' reflexivity and prolonged engagement, and triangulation of data (pp. 25-27). All these strategies were integrated in the current study. Firstly, member-checking strategy was used in data collection: a summary of the interview data was sent to each participant for their confirmation after the interview. Furthermore, the think-aloud processes were mutually explored in the stimulated recall session; and participants were asked to discuss among themselves the findings of the study in the focus groups at the end of data collection. Secondly, peer-reviewing strategy was applied in this study in two aspects: the transcripts of data were checked by another researcher and the coding

categories and interpretation of data were reviewed by the supervision panel; and aspects of this study have been submitted to journals for academic peer review. Thirdly, the strategy of reflexivity was also applied to this study both by self-reflexivity and methodological reflexivity (Hood, 2009). My position in the study was examined before the process of data collection (See Section 3.3.2). Moreover, as suggested by Borg (2001), a research journal (Appendix J) was kept throughout the process of data collection and memos in the process of data analysis, which was helpful for me to keep critical awareness of my positioning and the choices I made of research methods, strategies, and data. Finally, triangulation was used as a main strategy to increase validity (Duff, 2008; Mathison, 1988) in the study. Triangulation was originally used in quantitative studies in order to ensure accuracy of interpretation by confirming the consistency of findings (Mathison). However, Mathison argues that the value of triangulating data is:

> providing evidence - whether convergent, inconsistent, or contradictory - such that the researcher can construct good explanations of the social phenomena from which they arise. (p. 15)

Similarly, Duff (2007) argues that the value of triangulation is to reveal the complexity of reality rather than simply to seek convergence of data. Denzin (1984) listed four types of triangulation: multiple data sources, methodologies, investigators, and theories. In this study, multiple data sources and methodological triangulation were used. Data were collected by five data collection methods from tutors, which aimed to provide overall understanding of tutors' activity of assessment by collecting data on both beliefs and practices. Different sources of data were compared and triangulated constantly with each other. Aditionally, the data were collected across two teaching semesters so that there was time triangulation of data. By its nature, a doctoral research project has to be conducted by a single researcher, so effective investigator triangulation was not possible in the present study. As far as theoretical triangulation is concerned, the present study analysed data through a grounded theory approach rather than adopt an *a priori* theoretical position as a lens to examine data. (It may be noted that the use of activity theory was to illuminate the complexity of the findings, rather than analyse the data.)

The study also took consideration of the issue of transferability. Transferability in quantitative studies refers to the degree of replication of the study in other contexts. Stake (1980) notes that a single case in qualitative study cannot be generalized; however, the findings of the case can be 'relatable' (Bassey, 1981) to the readers' own experience so that "naturalistic generations" (Stake, p. 69), referring to meaning interpreted by readers who share similar contexts and experiences, can be achieved. Inspired by Stake, transferability was addressed in this study by the detailed description of the context of the study and the procedures of data collection and analysis, which is intended to provide implications for further studies in similar contexts.

With quantitative origin as transferability, the other two criteria, dependability and confirmability, refer to whether the study could be replicated or confirmed by other researchers; however, replication of data collection was not considered as this study was a case conducted in a natural rather than an experimental context. It would not be ethical to ask participants for extra time to replicate the process merely for the purposes of research. Moreover, the involvement of other researchers in interpretation of data would be problematic because of potential theoretical stand of each researcher. Therefore, the consideration of dependability and confirmabiltiy of this study included two aspects: technically, the data transcript accuracy was checked by another researcher, and the coding categories were examined by the supervision panel; methodologically, the two issues were addressed by the grounded approach, triangulation, as well as ethical considerations of the study which have been analysed in the previous sections of this chapter. The following sections will present detailed procedures of data collection and analysis.

3.6 Procedures of data collection

Data collection was carried out largely according to the research design, but with heuristic adaptations. Each method of data collection was piloted and modified according to the pilot results. The whole data collection lasted for a year (from October 2008 to October 2009). Preliminary analysis was made after each method of data collection. Data were constantly compared and contrasted at each stage to inform further data collection.

3.6.1 Pilot studies

In addition to the survey, a pilot study was conducted which involved six interviews, five think-aloud and four stimulated recall sessions, the procedures for each of which were previously rehearsed by audio- and video-recorded sessions with fellow PhD students. The interview, think-aloud and stimulated recall participants were tutors, most of whom were also PhD and Masters students at the university representing the possible range of my potential participants in real data collection. All participants in the pilot study had experience of marking written assignments.

The pilot interviews covered the following points: focus points for individual interviews, manner of interviewing, question types, probing skills, the digital recorder, and venues of interviews. The following points were piloted in think-aloud and stimulated recall sessions: the time span between interview and think-aloud and between think-aloud and stimulated recall, participants' instructions for think-aloud and stimulated recall, type of interaction between participants and the researcher during the two sessions, observation skills in think-aloud, stimuli used in stimulated recall. In addition, transcription skills were piloted.

The piloting confirmed that it was possible to collect data strictly following the ethical approval. However, in piloted think-aloud sessions, it was impossible to get students' consent in advance when the tutors were not sure whose assignment to mark at what time. The solution was that the name of the student whose written work was to be marked should not be known by the researcher, nor should the researcher request to keep a copy of the marked assignment.

The piloting was important for several reasons. Firstly, it confirmed the feasibility of the design of the whole process of data collection. Secondly, it revealed that the seemingly separate methods of data collection (interview, think-aloud and stimulated recall) were closely related to each other. They demonstrated an effective cycle of researching the process of teacher cognition. Finally, the piloting contributed to the preparation of detailed procedures of data collection. Other issues came under consideration, such as the recruitment of participants, the use of appropriate equipment, the setting, time span, and the appropriate manner of conducting data collection.

3.6.2 The survey

The survey was conducted in October and November of 2008 in the following steps. First, the questionnaire was published on the website of the faculty with a covering letter for participants. Then fifty-two tutors' e-mail addresses were obtained from the various administrators of the nine departments which employed tutors and each of these tutors was sent an email on 15 October, requesting them to answer the questionnaire via the designated web link. Twenty-eight anonymous responses were received from tutors in the nine departments.

Data were then grouped according to different departments. The data were analysed manually according to the survey questions. Comparisons were made between departments and a holistic view of the bio-data of tutors and their general attitudes of giving feedback was obtained. Fourteen participants provided comments at the end of the questionnaires. Content analysis of the comments was made. Issues of giving feedback revealed by the comments were listed for further exploration.

3.6.3 Participant recruitment for the following sessions

Tutor participants were mainly recruited by the survey and the 'friends of friends' introduction of colleagues. I kept contact with those survey respondents who had expressed interest in participating further in the research. It took time to establish rapport with potential participants and gain their trust which proved to be crucial in obtaining their consent to participate. The utmost effort was taken to recruit and maintain participants for the complete duration of the data collection. This included: more control given to the participants of how data were to be collected and how much time they would like to spend with the researcher; minimum interruption on participants' normal routine and schedule of work; and a friendly environment to promote communication between members of the community on the topics of their interest.

3.6.4 Formal data collection process

As planned, the data collection process went through two academic semesters in 2009. Altogether sixteen tutors were interviewed, nine of whom then participated

into think-aloud and stimulated recalls sessions, and seven of whom participated in the two focus group discussions.

3.6.4.1 Individual interview

The individual interviews were conducted according to the research design. Sixteen tutors volunteered to participate in the interviews. Participants were met individually and signed the information and consent form before each interview. The interview venues were chosen by participants: twelve interviews were at a café on campus, three at participants' office, and one at my office. The interviews lasted from twenty minutes to half an hour. The focal points were raised but flexibility was also given to participants who were willing to elaborate on relevant topics. All interviews were audio-recorded. The interview data was transcribed immediately afterwards. To enhance participant validation, a summary was sent to each participant for their confirmation, correction or amendment. Constant content analysis and comparison was done to inform further data collection.

3.6.4.2 Think-aloud sessions

Think-aloud sessions with individual participants were conducted at least three weeks after the interview to diminish the possible influence of the points raised during the previous interview on participants' actual practice.

Brief written instructions (Appendix F) were given to participants by e-mail in advance. Training was not given to participants because of the purpose of conducting think-aloud session as has been stated in Section 3.2.1.3. It also turned out to be practically unnecessary in this study: all participants confirmed their capability in thinking aloud. This was probably because of participants' work-related ability in verbalization and their familiarity with their work of providing feedback as well as their familiarity with the researcher and the study that had been developed from previous sessions of data collection and contact (Li & Farrell, in press).

Each think-aloud session was audio-taped. It began with greetings and brief small talk. The participant then introduced the assignment to me and I briefly explained the think-aloud process, making it clear that I wanted him or her to produce a

monologue and ignore, as far as possible, my presence (Appendix G). I assured the participant I would try not to interrupt their marking process at all.

Then I sat beside the participant and kept silent for most of the time, except showing my attentiveness by quiet back channelling or other short cues to motivate verbalization, although I did not provide any prompt to request their verbalization. It was evident that participants could not talk aloud at times when complex thinking was needed on how to write the feedback. However, no participant kept silent for more than three minutes before they started to talk again. All participants looked at the assignment rather than me. My main activity during the think-aloud session was to keep field notes which included the venue, time, type, and length of the assignments, operations of giving feedback, sounds (e.g., laughter) made by participants, the symbols they used as feedback, the in-text and overall feedback they gave (I could see or hear), the special features of the feedback, or decisions made by participants like erasing or changing their error correction or wording of feedback which were not talked aloud. While keeping field notes, I also marked the points that I would explore in the stimulated recall sessions.

3.6.4.3 Stimulated recall sessions

The written guidelines for the stimulated recall sessions were e-mailed to participants at the same time as the think-aloud instructions. After each think-aloud session, I asked my participant whether I could ask him or her some questions on the think-aloud process. Stimulated recall sessions were conducted immediately after the think-aloud session, except for the first case when my participant preferred to conduct a stimulated recall session the next day.

The questions for the stimulated recall sessions were the points I observed and highlighted in the field notes (Appendix H). The stimuli used in these sessions included the assessment feedback tutors had just provided on the written work and the marking schedules tutors used in assessment. At the end of each session, I asked participants for a copy of the marking guidelines or sample answers. Three of the nine tutors provided a copy or allowed me to photocopy the marking guidelines. One of the nine tutors provided me a sample answer for another paper rather than the one he used. He also provided me with the regulations for tutors in

his department and some other written documents. The length of the stimulated recall varied from ten to fifty minutes which largely depended on the time participants had after the think-aloud sessions and the type of assignments.

After the two sessions, I e-mailed participants a letter of thanks, asking them to confirm the summary of the think-aloud and stimulated recall sessions and clarify some issues that I came across when listening to the data. I received the confirmation of all participants, together with answers to my follow-up questions and their positive evaluation of the project.

3.6.4.4 Focus group discussions

At the end of the following semester, tutors who had participated in previous sessions were invited for focus group discussions. The focus group discussion was not only useful for the purpose of data collection, but also provided an opportunity for them to exchange ideas and co-construct understanding of the issues focused on in their work.

In the research context of this project, the participants of focus group were tutors of the same faculty, although not necessarily in the same department. Therefore, they were able to interact with each other on topics of interest they shared without the intervention of the researcher. Two focus group discussions were carried out: one focus group was composed of four tutors in the same department, and the other was composed of three tutors from three different departments. The focus group sessions lasted well beyond the time I had anticipated, indicating the participants were interested in the discussions.

3.7 Procedures of data analysis

The data analysis in this study took a grounded theory approach to the extent that categories of data were generated inductively, themes were sought in relation to research questions and the philosophical background of the study, and constant comparisons across data were made, initially through the application of NVivo *8*. It was also influenced by the six steps of thematic analysis suggested by Braun and Clarke (2006) in practical procedures, as discussed in Section 3.4. However, mainly the three key procedures of data analysis will be explained: organizing and transcribing data, initial coding for categories, and axial coding for themes.

3.7.1 Organizing and transcribing data

All audio-recorded data were initially transcribed by myself into NVivo8 because this programme offered the conveniences of storing sound and word files, transcribing, and analysing data in the same software (Bazeley, 2007). I set up the project in NVivo8 and created a folder called 'recording' which included three second-level folders: interview, think-aloud and stimulated recall, and focus group. I then imported the collected data into their respective folders, using the pseudonyms of participants as the names for the third-level folders.

One advantage of NVivo8 is that there is no need to transcribe every recorded word because the raw data of recording can also be coded and compared constantly. This was a clear advantage for this study because coding and analysis of each source of data were needed alongside the data collection process in order to provide information for next phase of data collection. About ninety per cent of the recording of each session was transcribed. The transcripts were organized into three main folders: interview, think-aloud and stimulated recall, and focus group. The time spans of silent periods were also noted down.

The process of transcription was time consuming. However, it was helpful for me to familiarize myself with the data and form an overall frame of the content of data. The transcripts were later checked by another researcher who was a native speaker of English in New Zealand. The checked transcripts mainly followed the transcription convention suggested by Du Bois (2006) (Table 1, p. xiii).

3.7.2 Open-coding for categories

The second step of data analysis was initial coding. This was also conducted in NVivo 8. I firstly coded one transcript of an interview conducted at the early stage of data collection. I read the transcript, highlighted any word, term or piece of sentence that summarily represented a complete piece of information which was used as a raw coding base. Then the raw coding bases were labelled by descriptive codes. In case a piece of information was illustrated by several sentences, I summarized the meaning with a descriptive code. For example, in the following piece of data, the highlighted words were used as raw coding base:

> I think yeah, I always like the **criteria**, but I'm just not as **heavy handed** as some of the other markers were. I tend to **mark a bit higher** because usually if it is a first essay, it kind of shatters **confidence** a little bit if they don't do really well. My students when they got their essays back they thought I'd **marked them quite easily** as well. They were very **happy** with what they got. (Mia, Interview)

These highlighted words were then labelled as the following initial codes: *attitudes towards criteria, higher marks, students' confidence, students' happy feeling*.

In the same way, each piece of data was open-coded. Then I re-read the data, grouped the open-codes into initial categories. By categories, I mean terms to show the hierarchical structures of codes, (called 'nodes' in NVivo8). For example, *students' confidence* and *students' happy feeling* were then categorized into a sub-category, *students' positive emotions* which belonged to a higher level category, *students' emotion*, which was under the broader category, *emotion*. In this way, data were condensed into summative terms and organized into tree codes or coding systems.

This open-coding process was time-consuming. The codes and categories were defined in a discursive way which included several rounds of reading and coding, comparison and adjustment. The benefit of the grounded approach was that all pieces of data were considered as a meaningful construct of the study so that the findings would not be filtered or limited by the research questions. As a result, emotion as an important category emerged from open coding, which was beyond the original intention of this study.

In addition, memos were also kept alongside the coding process. The memos were initially annotated in NVivo8, and were repeatedly consulted and augmented alongside the process of re-reading, making contrasts, comparisons of data, and seeking links between them. For example, the initial memo kept for the sample data above was: "Divergences between tutors in grading" (Mia, Interview Memo). At a later stage, "community of practice (lecturers and tutors)" was added to this memo, and when I sought links between categories, I noted down the following pieces of thought in the memo:

Grading as a tool to regulate students' emotion. Tutors used higher marks to generate positive emotions of students. Contradictions between tutors in grading. (Mia, interview Memo)

The think-aloud data were coded in the same way as the interview data except that I also compared the transcription with the field note of my observation, and added the nonverbal actions (such as erasing a word, going back to check the previous page, ticks, correcting a word) into the transcriptions to enrich the think-aloud data. Moreover, the think-aloud data was triangulated with stimulated recall data. Both think-aloud and stimulated recall data were open-coded, then categorized. The open-coding process allowed the opportunity for emerging codes and categories. As had happened in the coding process of interview data, codes such as *frustrating, happy, pleased* emerged in data of think-aloud and stimulated recall. These codes again were categorized as emotion. Finally, the open codes were clustered into four main categories: beliefs, practices, sources of beliefs, and emotion.

I used NVivo8 in these two stages of data analysis and benefited from it regarding organizing raw data and the detailed open codes, and viewing the coded data in a hierarchical structure. I also used its query functions in checking the coverage of a code within and across data. These functions also facilitated my comparisons between data in relation to a specific code and category. The following stages were then carried out manually to check the coding system, and more importantly, to seek connections between data for theoretical themes.

3.7.3 Axial coding approach for themes

I used the term axial coding mainly because I not only sought the relationship between categories but also the relationship between the coded data and research questions as well as the philosophical background of the study. This process included reviewing data and categories, comparing data based on coding and categories, reviewing literature, and using the findings to address the research questions.

Firstly, I manually reviewed the coding system in a deductive way: I applied the categories to the each set of data by reading the hard copies of data and

underlining the extracts for each category. The use of this process was intended to gain further understanding of the data and to check whether there were any data beyond the coding system. Meanwhile, I also examined the relations between the categories and sub-categories. The review finished with one change in the initial category: emotion was distinguished from other data and was set up as a main category because it represented a different direction from other topics of assessment such as grading and feedback.

Secondly, I compared tutors' beliefs with their practices. I also compared beliefs and practices among tutors. Convergences and divergences were found within beliefs and practices, and between beliefs and practices. This seemed to me a further step in understanding the categorized findings. Although at this stage the coding system had already demonstrated relations between data at a surface level, it could not explain the relations between the categorized findings and the philosophical underpinning of the study.

Therefore, I related the convergences and divergences to the philosophical framework that underpinned my study: between the literature of distributed cognition and tutors' convergent beliefs and practices, and between the notion of contradiction of activity theory and tutors' divergent beliefs and practices of assessment. Relationships were found between cognitive development and the contradictions and convergences in the activity system of assessment. Further analysis between the convergences and divergences in emotion revealed the relationship between convergent and divergent beliefs, practice, and emotion. This relationship was then re-examined within and across data. Finally, the theme was defined as interactions between emotions, beliefs, and practices due to convergences and divergences in the activity system of assessment.

3.8 Summary

By using a multi-approach of data collection, data were successfully collected from the tutor participants in the faculty. The collected data included 28 responses to the e-questionnaire, sixteen individual interviews, nine think-aloud and stimulated recall sessions, and two focus group discussions comprising seven participants in total. These data included tutors' beliefs about assessment feedback, their self-reported actual practice and on-going cognition of providing assessment

feedback, and their recalled practices. The data were analysed by a grounded theory approach. It was conducted by myself and was facilitated by computer software NVivo8. Using inductive open coding, I was not constrained by my research questions so that the important new category, emotion, emerged from data. Data were constantly compared and triangulated. Data of beliefs and practice complement each other. The main convergences between all sources of data were the interactions between emotion, cognition, and action. The quality of data collection and analysis was increased by triangulation. The methods of data collection and analysis fit the research focus. More importantly, the inductive approach of data analysis provided the opportunity for the emergence of the contextual factors that affect tutors' cognitive development.

The key findings will be presented in the next chapter. It will begin with a statement of the research questions that guide the present study. These research questions are intended to address the gaps discussed above.

CHAPTER FOUR: FINDINGS

4.0 Introduction

This research project aims to contribute to the theoretical framework of teacher cognition by exploring the convergence and divergence of the beliefs and practices of subject tutors in the activity of giving feedback, both as individuals and as members of specific communities of practice in a New Zealand university. The research questions that the present study aims to answer are as follow:

- --What do subject tutors in the specific context believe about giving written feedback on students' assignments?

- --What are their actual practices when giving feedback?

- --To what extent do their actual practices converge with or diverge from their beliefs, individually and across departments/disciplines?

- --What are the socio-cultural factors that influence tutors' beliefs and practices?

- --How do the findings to these questions add to academic understanding of what constitutes teachers' beliefs, and the possible tension that arises in putting these beliefs into practice?

Data were collected chronologically across an academic year (from Oct. 2008 to Oct. 2009) by five methods: survey, individual interview, think-aloud, stimulated recall, and focus group discussion. The audio-recorded data were transcribed and stored in NVivo8. The transcripts were checked by a fellow researcher. Data were firstly open-coded and axial coded in NVivo8, where categories were set up. Constant comparisons were made among the same source of data and between different sources of data. Convergences and divergences of beliefs within and between individuals were examined. The beliefs were compared with the practices. Convergences and divergences were found between beliefs and practice. Then data were analysed again manually by applying the categories to all data as a double check of the validity of coding and to gain deeper understanding of data. Samples of coding and categories were also checked by the supervisors. The core categories were then selected for themes.

The findings enabled me to address the original research questions. Furthermore, two important findings emerging from the data provided the opportunity to explore the research questions in more depth. Firstly, it was found that tutors' beliefs and practices were influenced by the context of assessment activity. More unexpectedly, arising from data collected in interviews, think-aloud and stimulated recall sessions and the two focus group discussions, emotional issues were observed to be a significant aspect of assessment activity. The key findings will be reported in this chapter in the following sequence: 4.1) Beliefs about assessment; 4.2) Practices of assessment; 4.3) Sources of beliefs; 4.4) Emotions and the activity of assessment; 4.5) Summary of the findings.

4.1 Beliefs of assessment and feedback

Tutors' attitudes towards assessment and feedback were first elicited from the open-ended questionnaire items, but their beliefs were more deeply explored in the interviews and focus group discussions.

This section will report the key findings of beliefs. The structure of this section is 4.1.1 Good written work; 4.1.2 Feedback; 4.1.3 Grading; 4.1.4 Beliefs about criteria; 4.1.5 Summary.

4.1.1 Good written work

Beliefs about what constituted good written work were found in data of the interview and focus group discussion. As is shown in Figure 4.1, altogether sixteen aspects were mentioned by interview participants.
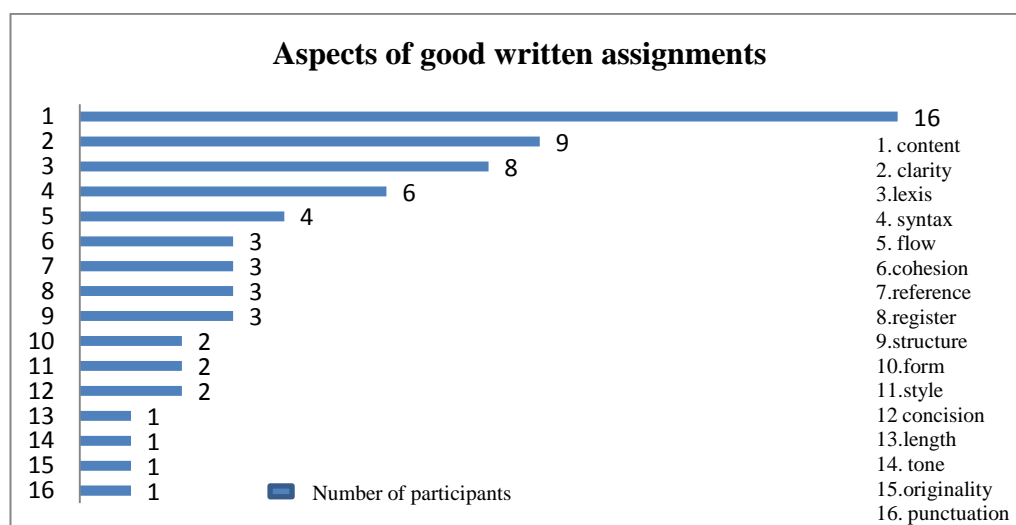
Figure 4.1: Beliefs about good written work held by interview participants

All sixteen interviewees believed a good written assignment should focus on the content, by addressing the topic or answering the questions, providing evidence or examples. For example:

> Researcher: So what is a kind of good assignment in your perspective?
> Cecile: One who has followed the guides and answered the questions. Yeah basically they have done what is asked of them. So I think it is more important to have the right content, than it is, I mean, obviously grammar and spelling are things that are important but it is more important to have right answers, and have the right content than … (Cecile, Interview)

> George: I think they are good because they have clearly and concisely and more importantly they have been able to relate it to the literature, been able to get the topics that we asked about them. They have been able to write about those topics. (George, Interview)

The next important aspect, clarity of expression, was mentioned by nine participants (Helen, Simon, Mia, Mark, George, Emma, Jan, Henry, Frank).

> Researcher: What is a good report?
> Mia: I don't know. It always changes for each person. I just tell everyone to make it clear and concise. Don't write anything extra than you need to because you are just going to go over the same things over and over and kind of muddle yourself.

> Researcher: Then what is a good essay for your students?
> Mia: I tell them the same thing. Just be clear and back everything, back up your arguments with evidence, and make sure you refer to whatever the question is.
>   …
> Researcher: Why do you think clarity is so important?
> Mia: Because I have got so much trouble with not clear with my own essay, but also if you cannot understand the point you read in an essay, it is quite hard to judge it if it does not make a lot of sense. (Mia, Interview)

Clarity was also discussed by tutors in a focus group:

> I mean there's all these different theories, what, what sort of ones do you think are particularly you know amenable to possible explanations for this seemingly /// decline in this skill of being able to write clearly or construct coherent arguments and #that #sort #of #stuff. ## there's the change in teaching ##, that's what I remember lecturers saying to me. (Henry, Focus group A)

The third important aspect of good writing was related to errors at the lexical level such as spelling, terminology, and word choice, and was mentioned by eight participants (Helen, Simon, Mia, Mark, Cecile, Eva, Henry, Jan). The fourth aspect was at syntax level, mentioned by six participants (Helen, Mia, George, Jo, Henry, Jan). Less mentioned aspects were flow or fluency (Helen, Eva, Jan, Martin), connection (Anna, Simon, Eva), referencing (Helen, Anna, Simon), formal language (Anna, Mia, Emma), structure (Emma, Eva, Simon), form (Martin, Henry), style (Eva, Henry), and concision (Mia, George). Other aspects of language were each mentioned by one participant: tone (Simon), word count (Mark), punctuation (Helen), and originality (Frank).

The most detailed description of good writing was given by Simon:

> Researcher: So can you summarize what is good writing?
> Simon: Err, may be on a general level, it is clear, it is structured, it has, err, a clear connection between what's being said and what conclusions are being drawn from that, you know. It's well referenced, recognizing where you get the information from. So that is generally good writing, it has a structure, and has, you know, arguments and conclusions, err specifically, it would be things like having the appropriate tone and voice, so you are using, you know, depending on the essay, you are using an objective voice, and you wouldn't be saying I think this, you would be saying, you know, Miller (2007). /// There is lots of little phrases that you've got to familiarize yourself with / vocabulary of kinds of ways of structuring sentences. /// So-and-so argues that, or in this book, so-and-so presents this idea. There is lots of little ways of knowing how you can get your information out in an academic style. That once you have them, they come naturally and it just makes writing so much easier. (Simon, Interview)

He also clarified the main difference between good writing and a good assignment:

> Researcher: Do you think good writing is different from good assignment? Or do you think good writing is good assignment?
> Simon: No. I think good writing is something entirely different from a good assignment because it can be really insightful and have lots of merits and know the topic, show some real insightful understanding of the topic but /// you can get a good assignment with bad writing so long as your writing is clear. (Simon, Interview)

However, most of the interviewees found it difficult to generalize about good writing, and needed to focus on the specific assignments they were allocated - as can be seen in the following two examples:

Researcher: So what is a good piece of writing for your students?
Anna: I wouldn't really know. What you mean in terms of what they could look at or ///
Researcher: Say, this is a good assignment.
Anna: Well they've got the topic, and then they are presenting the ideas and they are actually returning to the topic and they are actually evaluating it in relation to that, and they are using the recommended text, and they are thinking about it and they have covered what the assignment is meant to be about I suppose, that's good writing. And I mean it isn't necessarily about, I suppose every assignment is different. It's not about their own point of views a lot of it's about being able to encapsulate other people's ideas around the subject which is what undergrad University is about really, learning about the discipline and who's involved in the discipline and new ideas, new concepts. (Anna, Interview)

Researcher: What is the standard of good writing for [[department name]] students?
Helen: Err, I am not sure really. No, I think it varies. Is this exchange necessary?
Researcher: Say if you mark assignments, how can you say this is good?
Helen: They need to present their ideas. They need to give evidence for their ideas. They need to give examples or evidence or whatever, and they need to reference those examples or stuff that they've got from other books and stuff accurately, that's very important. (Helen, Interview)

In focus group discussion, Helen talked further about reference issues:

And you know, especially the students that I taught in the first semester, right? And we spent AGES on referencing. And then I got some of them in the second semester, and some of them had got it and some of them hadn't, and I was thinking 'I'm a failure', but I was also thinking /// they know it. (Helen, Focus group B)

Both Helen and Anna had a better understanding of a good piece of assignment they marked than a good piece of writing in general. Both emphasized the presentation and discussion of ideas as an important feature of a good piece of written work (assignment).

For Mark, a good written assignment leading to a high mark was a demonstration of the understanding of terminology by giving examples.

Researcher: What is a good assignment?
Mark: For this paper that I am tutoring I think, I encourage students to, basically, well obviously good literacy, a good use of, a good fistful of the terms and the concepts that the lecturer

has taught them and everything. Without them, like I said they are the key ingredients of getting a really good mark. And without them you won't get that good mark that you really want you see. So I think that is the key thing I think really. I think if you use those terms, and you don't have the accurate good literacy you'd still get to pass, but if you use those terms, and also have good literacy and you show good understanding of those terms and give examples as well on top of that, you will basically get a really good mark. That's what I think personally. And I think it is important to give examples because it shows that you understand it as well, you know. Some students I was marking they tend not to give examples. That shows to me that they understand it but they cannot give examples with it and show even more understanding, depth, you know. The more examples you give with those terms and concepts, the better. (Mark, Interview)

Like Mark, George also related a good assignment to a high grade:

Researcher: Can you summarize what is a good assignment?
George: An A+. There should always be something wrong. I am not God. I don't do perfect. I leave that up to other people who think they're God. I do, yes, if I give an A+ I think they are good because they have clearly and concisely, and more importantly, they have been able to relate it to the literature, been able to get the topics that we asked about them. They have been able to actually write about those topics. (George, Interview)

4.1.2 Feedback

Three key findings of tutors' beliefs about feedback were found in the data from questionnaires, interviews and the two focus groups: the purposes, the foci and the strategies of giving feedback.

4.1.2.1 Purposes of giving feedback

Fourteen of the 52 survey participants responded to the open-ended question: Please feel free to add any further relevant comments regarding feedback on students' assignments. The most common comment (5 participants) was about the purpose of feedback. Four participants believed that feedback aimed to help students make improvements. For example:

I mark under the assumption that students will read feedback (I hope this is the case) and actively try and improve regardless of ability. (No 3, Open-ended comment)

My ethos is to help them as much as possible. All too often the assumption from markers or tutors is that the mistakes students

make are obvious, or due to a lack of editing. While there is certainly truth to the lack of editing, the fact is that a lot of students (and even some of us graduates) need constant feedback to progress. (No. 14, Open-ended comment)

Similarly to the open-ended comments in the questionnaire, all interviewees believed feedback was given for the purpose of making improvements. For example:

> Researcher: So what's the purpose of giving feedback?
> Emma: For improvement. It is not to say this is what you've done wrong. It is so that people can improve on those things. That's why you don't really give feedback on exam because by that point it is too late. (Emma, Interview)

Consistent with his notion of a good assignment, Mark believed improvement included meriting a higher grade:

> If I put a comment on it, and I told them to look at this and look at lecture three, they will look at it then and then they will have a better, clearer indication what to do and then the following week, and then the following weeks, then in the semester, they will understand more better and then hopefully they get a better mark in their exams as well. (Mark, Interview)

In addition to helping the students to obtain a higher grade, George expected more improvement:

> Researcher: So what's the purpose of giving feedback to your students?
> George: To improve their grades, there's no other real thing. It'll improve their grades and everything kind of flows off that, things like improving their writing, improving their intellect, improving content //// and to also improve themselves as people. (George, Interview)

However, one questionnaire respondent thought it was mainly for the purpose of justification of the grade:

> I think that it is most important to give feedback when marking assignments. I make comments as a safeguard for me so that it is clear to the student where marks are being taken off. I also write positive comments so that the student knows where they have gained marks! (No. 12, Open-ended comment)

This purpose was also expressed by two interview participants:

> I prefer written feedback that explains how and why you got the grade. (Simon, Interview)

> I really wanted to be helpful and you know so that the students would be able to know exactly what it was they had done right or wrong and why I gave them that mark. So I wanted to justify it for myself as well. (Maria, Interview)

On the other hand, most interviewees believed they could not identify much room for students' improvement when marking well-written assignments. For example, Helen found "For some people, it is very easy to find the positive things and quite difficult to find things they needed to work on". Mia believed if the assignment was well written, she would not "bother to comment". Emma believed "Feedback could be short for some students who you cannot find many things to improve". Cecile said she would only say "excellent work" if the assignment were really good. This point was supported by Simon in a focus group discussion:

> If it's a really good essay you can just go tick, tick, tick and give them an A. (Simon, Focus group A)

In focus group discussion B, the tutors discussed how the students did not make significant improvement over years on some aspects of writing:

> @Yeah, I know but, but, they know this. They know what they have to do. We've been through it so many times. [mm] You know, so yeah, you know, I think there's, there's a sense that in the first semester, you kind of, and quite rightly don't expect – they don't get taught how to reference, especially [[subject name]] referencing, you know, it's different from most of the rest of the university. So they don't know it. So they, – you know, you do give them a bit of leeway, but when you know they HAVE been through that first semester and they've had ample opportunity to learn it, and then you're still teaching it to them in the second semester tutorial, and they're still not getting it, well they're being lazy with it. (Helen, Focus group B)

Overall, it was not clear whether the purpose of improvement could be achieved. According to the survey, all except one of the respondents believed their written feedback was effective. However, over 40 per cent thought students would not read feedback carefully, which put in doubt the assumed effect of feedback. When she was asked in her interview whether she believed students would read her feedback, Maria said that students "don't read it anyway." Mary expressed the same doubt:

> Researcher: Do you think they read your feedback?
> Mary: I doubt it. I mean when I did this paper, all I wanted to know was 'What did I get'. You know, but I, I don't know if they do anyway. (Mary, Interview)

This opinion is similar to that of Jan. Moreover, it seems that students' attitudes towards feedback also affected the motivation of the tutors to provide feedback:

> Researcher: Do you think students read the written feedback?
> Jan: I often think they don't. And that is one of the reasons why I don't want my first year tutors to write a big load of stuff. I mean we cannot really afford to pay them if the students aren't going to read it. And certainly I mean at the end, there are always ten per cent of the assignments never even picked up. (Jan, Interview)

Other interviewed tutors believed students varied in their attitude toward feedback: Eva 'hoped' and Frank 'thought' that students would read their feedback. Cecile went into the matter in some depth:

> Researcher: Do you think students will read the feedback?
> Cecile: I think some students probably might have a quick look at it. But most of them probably don't pay attention but there is ones who are. If they haven't improved over the next assignments you know that they are not really paying attention or if they keep making the same mistakes they are obviously not paying attention to the feedback. (Cecile, Interview)

George held a slightly different view:

> Researcher: Do you think your students read your feedback carefully?
> George: I know they do. I don't know if they all do but I know some of them do. (George, Interview)

Divergent opinions were also found in focus group discussions:

> Henry: Does anybody actually get feedback about the feedback? Do you actually think students read the comments on their essays and take note of them and everything?
> Mia: Yeah, I've got a couple that came back with like their essay with comments on them and like 'Can you explain this to me a bit more?' (Focus group A)

To sum up, findings from survey, interview and focus group demonstrated that most tutors believed the purpose of giving feedback was to help students improve; however, they were rarely explicit about what it meant by improvement other than getting better grades and making fewer grammatical errors. Moreover, these tutors were not sure whether all students would read their feedback and make improvements.

4.1.2.2 Foci of feedback

Data from the survey, interviews, and focus groups strongly indicated that tutors focused not only on content in assessment but also paid attention to formal issues. The survey data demonstrated over 89 per cent of tutors focused on the content of written assignment, but more than 64 per cent also attended to grammatical errors and organization at text level.

All sixteen interviewees believed they focused on content while assessing students' written work. For example:

> Researcher: So you focus on the content.
> George: Yes. I was not after structure, I was not after /// The reason I didn't look after the structure is that it was not an essay it was only a response. A paragraph of response. (George, Interview)

> Researcher: So you focus more on concepts or content.
> Mia: Yeah, that is what the other markers told me to do, rather than being lenient. (Mia, Interview)

However, like the survey respondents, the interviewed tutors had also to pay attention to formal issues in order to assess the content:

> Researcher: So, you focus on content or ideas.
> Helen: Yes, very much. I mean form is important you know that they write that they don't make too many mistakes. That they proofread their work and /// (Helen, Interview)

> Henry: For me the, the main thing is the ideas, [oh]. Um, or that is what I think is the most important, [mm] so the students kind of understand the sort of key ideas about the the – what the course is? [mm] but just, you know [[name of the general field]], and / [yeah] So um, // yeah so often marking it's sort of like 'this this person I can kind of get what they're saying [mm] and I think I DO understand what a lot of their ideas are', but just as a reader, it's hard / for me // to really, I- because it's not clear? in their writing? So I kind of have to / assume that that's what they're about sometimes? / [mm] and that's possibly the problem. (Henry, Interview)

As pointed out by Henry, the formal issues affect their understanding of the content of students' written work. In order to assess the written work, Anna sometimes had to rewrite the sentences:

> Researcher: You have to rewrite them?

Anna: Sometimes you can't understand, you try to give them marks for content but their actual writing skills may not be up to scratch. (Anna, Interview)

The formal issue in the assessment was also supported by data from both focus group discussions:

Simon: If it's a really good essay you can just go tick, tick, tick and give them an A, but / if / you're trying to structure an argument and they just don't have any structure there / sometimes you know I find I have to rewrite an entire paragraph for them, and say, you know // – and then do I mark them on what I think they're trying to say or what they really wrote.
Frank: Yeah. (Focus group A)

This issue was also discussed in focus group B discussion, with more information on formal feedback to different written work and students' backgrounds:

George: Yeah because I did essays as well. [mm] Yeah, um, yeah, correction's an interesting one with me. Er in the labs because the labs are hand written, I tend not to worry about style and grammar? [oh okay] because they're just like answering questions in short sentences and they've only got two hours to do it. So it's very quick that they have to do it. And um, some of them / like the internationals, I'm way way more lenient on, and I've been told off before for that, [mhm] but um to me, if you've just come from another country that doesn't speak English and doesn't write in English, and you get thrown into a place where that's all that anybody knows, / um / you're going to struggle, [mm] for a little while at first. And I think that / like by, well in the lab anyway, I tend not to worry about that. If it was typed, if somebody types instead of print, then that is completely different. You shouldn't hand in something that's been typed with a spelling mistake, [mm] because Word will check for you@@
Helen: But then you get the 'forms' and 'froms' and the whole, and I've actually been pulled over because they annoy me. Um < ... > yeah, they do annoy me. We'll be honest here, people who over ###, I have that feeling anyway, I had an English degree to start off with, you know, but also I think people rely too much on spell-check.
George: Yeah.
Helen: You know, and they don't actually read it through afterwards to check it actually makes sense, / you know, because spell-check won't pick up / won't necessarily pick up those sorts of spelling mistakes [yeah], they'll only pick up a word / that they can't recognize.
George: The one that gets me that I write on the bottom of the essay is when everything is spelt right, but it's spelt right for American spelling?
Helen: Oh okay.

George: It's like 'characterizzze' @@@

Helen: I tend to underline, circle, even for students that might not speak English.

Maria: I thought maybe mine's not going to be good, I'm not trained to do this, I don't know when their syntax is not going to be good. But it was so obviously ##, I was like 'No, that's not right, what were they thinking?'

George: I love those ones where you read a paragraph and you think 'What the hell was that all about? What was the point of THAT?'

Maria: Yeah, like I can't even see what they were TRYing to get at.

George: The ones that throw me out are the paragraphs that are about three pages long with no break?

Helen: I wonder what they teach them in school, actually. (Focus group B)

These findings revealed that the formal issue was the major barrier to tutors' assessment of the content, which explains why tutors believed clarity was a main feature of good written work (c.f. Section 4.1.1). In fact, all interview participants provided feedback on errors (c.f. Section 4.2.2.2). Therefore, it can be understood that the primary focus of assessment was intended to be content, but feedback included both content and formal aspects. For example:

Mia: Sometimes I'll underline a paragraph and I'll write on the side or underline a sentence and say this doesn't quite make sense you might think of rewording it. Or I'll mark out a section on the border and say this is not quite right. But otherwise if it is good I just put a tick next to it, that they have good ideas, and everything is fine.

Researcher: And then you will write a comment?

Mia: Usually if it is good I won't bother with a comment. And if it is really good, then I'll be like 'this is a really good point' and 'I'm glad you put it in.'

Researcher: If it is a bad one you may write something at the end?

Mia: Usually, I'll do just a one or two sentences saying 'This is where you could improve, this is what I liked about.' (Mia, Interview)

4.1.2.3 Positive/negative feedback

All participants of interview and focus group discussion believed feedback should be positive. For example:

Researcher: Do you give them written feedback?

Jo: Yes, we have. Not every time, but it is when we are instructed to give some feedback. Last week's lab I was able to say you know 'Good work John,' or 'Well done' or 'You did this very well' but this week, er, the lab was out of a number. So the

total was out of seventy, and so they got fifty out of seventy, or thirty out of seventy.

…

Researcher: You tend to give them some positive feedback.

Jo: Yes, positive, mostly. Unless they do a terrible job and then I have to say but it is often not the case. (Jo, Interview)

Cecile: So I always make sure I gave some positive feedback. And point out things that were good. And also I try to give examples of things. Not every time I do it, but most of times I'll say you could have done more here for example blah blah and then write out what they could have said… (Cecile, Interview)

Tutors believed that positive feedback can be more acceptable for students and effective in achieving the goal of improvement. For example:

Researcher: So what is your purpose of giving feedback?

Helen: Yeah I think probably to help them so they can succeed, so that they can get better ///

Researcher: To see improvement?

Helen: Yeah. That is why I like the idea of giving them positive comments /// Say 'You have done this very well' and then a constructive thing to help, not 'You didn't do this so or I think this was hopeless.' (Helen, Interview)

However, it seemed that there were different understandings of what positive and negative feedback was. Six interviewees (Mia, Mark, George, Jo, Cecile, and Emma) described their understandings of negative feedback, which was based largely on their own (previous) experience as students receiving feedback. For example, Mark believed negative feedback was the sort of feedback that unfairly assessed his ability with a low mark:

… The comments that the lecturer gave me which I thought was a bit, you know, a bit offensive, I thought it was like he wrote down 'Not good enough, your English isn't great and you need to learn, have classes on English, and everything.' And I thought it was quite, you know, I thought that was quite offensive. What really ticked me off was that I had two essays come back at the same time that was an only B- I had, but the other essay I had an A+ saying well written and everything like that. I know that I didn't put a lot of effort into that one, but for a lecturer to write that bad, I thought that if it was an international student, international students could get hurt by that I think. And yeah, I thought, and I wouldn't generally share that, you know, I wouldn't, say, the lecturer told me to have English lessons. You know, I thought it was quite bad, you know. I tend not to give those negative comments. (Mark, Interview)

Cecile believed negative feedback included harsh words:

> One of my supervisors could be quite harsh, and would just write 'Yuk' next to the sentences or paragraphs that he didn't like. So I thought that was a bit harsh, so I think it probably influenced me to be more positive and try to put more positive type of feedback. (Cecile, Interview)

Mia and Helen also believed positive feedback referred to positive expression of negative aspects:

> I think I've always found it easier when I got positive feedback, saying where I can improve rather than saying you did this wrong. I always trusted that more. So I always found that more appealing. (Mia, Interview)

This belief was also expressed in the focus groups discussion. For example:

> Mia: Do any of you give like, real negative feedback?
> Frank: I try not to, sometimes I have to, I just say – well usually I say /// 'This is what I liked about your essay, these are the ideas I thought you're on the ball with, BUT ///
> Mia: You missed the point.
> Frank: Yeah @ there are a lot of things you still need to work on. But I try and be as encouraging as possible, I guess.
> Henry: Constructive criticism.
> Frank: Yeah, constructive, yeah.
> Henry: Key phrase I suppose. (Focus group A)

George believed the number of negative aspects referred to by the marker may also make feedback negative, besides the choice of words:

> …because all my lecturers pointed out good and bad points. They never ever pointed out all negative points. They never said all these things are bad and then full stop. They always used words like 'This could improve your grade.' They always made very a subtle, you know, it was there, it was obvious (expression), but it wasn't saying you are really bad at this. (George, Interview)

Thinking about her present role as tutor, Jo seemed to believe negative feedback referred to negative aspects of the assignments, which made it difficult for her to give feedback:

> Maybe the only difficulty would be that I don't want to be too negative. I want to be positive even if they have got it wrong; I don't feel like I should be the one that says, 'Hey, you've got it completely wrong. You don't know what you are talking about.' (Jo, Interview)

Anna mentioned that being positive was a technique she used in giving feedback. Positive feedback could be positive 'slogans', by which she meant positive comments referring to no specific points (e.g. Well done.).

> Researcher: Do you use any techniques or strategies in giving feedback to students, in writing the comments?
> Anna: There's a whole lot of generic positive slogans that you have on hand and then you personalize this for that student. But I mean there is only so many ways you can go 'This is a very good essay, you have done well.'
> Researcher: Where does that slogan come from?
> Anna: Just from our cultural understanding of what positive feedback is, I suppose. And the way that you write something like you wouldn't, I think there is a type of language that is used and that it is not like 'Hey, man this is really good.' It is like this was a good essay because you have, and might refer to the title, you have done the things were asked, you've demonstrated you have, em, you know depending on what subject is you have demonstrated, and understanding, you have drawn on the text that were recommended, and have written it in a coherent way, you know, so I mean you actually referring it back to what the expectation and goals are of the assignment. (Anna, Interview)

It seemed that the positive slogans mentioned by Anna were a common technique used by tutors to be positive. In addition, tutors also believed ticks were also part of positive feedback. Actually, the use of tick was an instruction Jan gave to the tutors in her department:

> I mean you can't really do more than say 'Walk in any time, ring me up, email me' you know, I mean I cannot really do any more than that. So I think students who are on the receiving end of assessment quite often misunderstand the motives of the markers /// But students have said to me in the past, that every time there is something written on their essay, they feel they have been corrected. And I think that is very negative. And I don't think that is the motive of the lecturers. I mean some of it is positive, I try and encourage my tutors to put lots of ticks if they like things, you know, tick, tick, tick, if it's very good /// and I try and encourage my tutors not to put too many crosses. You know, wrong, or crossing things out. But I have different standards for different people. I mean those tutors, I have to train them every year to do what I want. (Jan, Interview)

The use of ticks was also mentioned in focus group A and was related to time issues:

> Henry: I probably myself agree with this /// I tend to /// for positive comments /// write, you know, just /// one word or just a

few words, whereas if you're trying to give some constructive feedback about some problem, that's much more in-depth.

Mia: Like you ####

Henry: Yeah yeah yeah. I thought, if you've got a stack to mark, it is just a 'This is real good' or 'Excellent work', or some /// little // phrase there that you just sort of use for the good stuff.

Frank: Yeah.

Henry: But often embellish – the comment at the end is where you know give them more, more fulsome praise if they've done a good job, so you know get into detail about it.

Frank: Mm, nah, definitely.

Simon: I found myself thinking back to the marks I got, what people wrote on my essays and that sort of stuff and, and one that they always 'good effort' but when I'm thinking of stuff to write on good essays you know, I think about what maybe [[name]] wrote on my papers, [right] all that kind of stuff, back when – yeah

Mia: ##### or something.

Frank: And I always get worried that I'm not giving enough feedback, because I mean if it's a really good essay, you know I just go through each paragraph tick, maybe put, you know, 'good', you know, 'great points', that's like what else can you do really, but, yeah. (Focus group A)

Moreover, a high mark was believed as being part of positive feedback:

Researcher: So why do you tend to mark them easily?

Mia: I think the students learn a bit more form positive feedback. (Mia, Interview)

All sixteen participants believed good feedback should be positive but should include both good and bad aspects of the assignments. For example:

There is a mark and then a written comment, which is either to reinforce what they have done, and to actually point out what they haven't. Positive reinforcement. (Anna, Interview)

Just basically things they have done well and things they can improve on. (Emma, Interview)

This point was also mentioned in a focus group discussion:

Helen: So you guys provide a lot of feedback?

Maria: I do, yeah.

George: I do.

Helen: Positive AND negative?

Maria: I always start with positive and work to negative … (Focus group B)

4.1.2.4 Summary

The first key finding was that tutors believed the purpose of giving feedback was to help students improve, although they were not sure whether this purpose could be achieved by their feedback. Some tutors mentioned feedback was also used to justify the grade they awarded. The second key finding was about the foci of feedback: all tutors focused on content but also paid attention to formal issues because it affected their understanding of the content. Tutors also believed feedback should include both good and bad aspects of the written work. The third key finding was that tutors believed feedback should be positive. However, they had various beliefs about what positive or negative feedback was. The most common techniques for being positive included using praise or encouraging words, ticks, pointing out positive aspects, and grading up.

4.1.3 Grading

The issue of grading was found to be the major concern of tutors in all the phases of data collection. In focus group A, tutors spent about 50 minutes in discussion of the eight focused topics, and about seven minutes on this topic - longer than the time they spent on other topic. In focus group B, tutors spent about 52 minutes discussing the eight topics, among which seven minutes were spent on marking. Then, when talking about feedback, the conversation shifted to grading for almost five minutes. Grading was related to other topics as well. Altogether, more than ten minutes were spent on the topic of grading which seemed to be the major concern for participants.

Tutors took various measures to keep consistency of grading such as comparing students' work, and exchanging opinions with each other or with the lecturers. However, inconsistency was unavoidable because tutors had different beliefs about what should be marked down. The key findings of grading will be reported in this section in the following sequence: consistency and allocating grades.

4.1.3.1 Consistency

Tutors were all aware of the importance of providing a fair grade, to provide which, they believed they should maintain consistency in their marking. For example:

We need to be consistent. If we are not consistent then we can't really grade people /// If my class was all getting A+s because I think, oh well! they're all very nice people, I'll give them A+s and if it was like, I hate Mike, I'm giving him all Ds. We are not actually marking their actual work, are we? We're marking what we think of them. (And we should never mark a person based on person?) Because that's judging people. And we really shouldn't do that //// But I think that if you know a person is trying hard and is really working at it, I think that you will end to be slightly easier on them. (George, Interview)

In order to maintain consistency among a cohort of students, tutors generally made comparisons between the marks they gave to different assignments. For example:

Researcher: You don't compare them?
Eva: For grading I will. But I think I'll do as Emma did and go through and just mark them all first with a pencil and write comments and then at the end, when I kind of know what I've put for them all, then I can go and compare. (Eva, Interview)

Similarly, Anna believed assessment was based not only on each assignment but also comparisons "across the whole group as well" (Anna, Interview). This usually took several rounds of reading:

It usually takes me quite a while, er, I like to read through an assignment, er, at least twice before I can make a decision of what the grade is because I find that you read an assignment and think 'Oh yeah! That's a B kind of thing.' And then you come back to it later and read it again you'll think 'Oh no! Actually they've got good ideas in there, I'll give them you know a B+ for it.' (Frank, Interview)

In focus group discussions, tutors mentioned the need to read the assignment several times for different aspects of assessment:

Helen: I probably marked them all three times – or probably read them all three times.
George: I read my lot twice. When I was doing the essays I'd read them once and have no marks on the piece of paper, and then I'd go back and read it again, scribble, scribble, scribble ///
Helen: See I did it the opposite way, [right] I'd go through and I'd mark them for grammar and /// and the spelling and stuff like that, and then leave the comment and the grade /// (Focus group B)

Tutors were also expected, and tried, to be consistent with the lecturers or the senior tutors and other tutors if they marked the same assignments. All tutors received oral instructions and written guidelines on marking. Also, assignments

marked by tutors might be checked by lecturers to ensure consistency. Some of the tutors specifically mentioned their experience of re-marking the assignments to be consistent with the lecturer. For example:

> And then after I'd given some of them back to [[lecturer name]] and he said that I was being a little bit hard, so then I went through them all again to see whether I needed to increase their marks, whether, based on what he'd told me, whether I needed to increase their mark or whether I thought it was still consistent with what I'd done with the other ones. I found it quite a complicated thing. (Helen, Interview)

Various measures were also taken to maintain consistency of marking among tutors. The tutors in the same department (George, Jan, Jo, and Mark) mentioned they had a meeting about tutorials and marking every week.

> Every Monday we have a meeting at 9 o'clock. And that is where we go through the lab, and what we need to know about the lab. (Mary, Interview)

A similar practice was mentioned by the tutors in other departments:

> And there's always a meeting, / [mm] with the tutors and the lecturer o- or the course supervisor, just to go over – everybody reads over it and then is there any questions about what does this mean, just to clarify everything, with ALL the tutors, [ohh] so everybody is obviously on the same // has the same approach, if you like. (Henry, Interview)

Emma mentioned that she exchanged opinions with the other tutor for the same paper. Cecile mentioned they used "cross marking" to ensure consistency among tutors to give "enough and the right feedback". Mia mentioned she had the experience of marking the same assignment with other tutors to ensure the same level of marking:

> The class I tutored last year, there were three of us who met weekly and when we were marking, we traded essays, to see if we were all on the same page with our marking, we were all roughly about the same, but it was interesting that we had all come from English backgrounds, so a lot of the essays we were picking up on the grammar mistakes, and the spelling, as well as the concepts. But we found we had all marked around the same level, although they were a bit more harsh with one of the essays that had to be failed because I had given them slightly higher marks. (Mia, Interview)

In both focus group discussions, tutors also talked about the strategies they used to be consistent when grading assignments; similarly to the finding of the interview,

the most commonly used strategy was checking the previously marked assignments, and then perhaps making some adjustment:

> Mia: Do you go back and do some of the first ones again after you have finished, just to make sure they are on the same page?
> Frank: Yeah, oh usually I mark them all than, and then, just quickly run through them all again just to – 'cause sometimes I'll give someone a B and then read it again and just think 'Oh no! They actually had some good points' or /// 'Oh no, they actually didn't do very well' so ///
> Mia: It doesn't compare with this one that I gave a B or whatever //
> Frank: Yeah, exactly it. (Focus group A)

> George: Did you ever, um, well like, I did this um, with the [[course name]] like I'd marked 38, cause each tutor marks their own lab group, and then I'd gone back after I had marked them and /// sort of changed some of the marks, so that they would be /// you know, in a more stratified thing – cause we got told they had to be /// like more stratified than ///
> Helen: Yeah? //// em.
> George: So like such //
> Helen: [Yeah I know what you're saying.]
> George: [In comparison with the other] people you know, [yeah], so like this person here wrote this like fantastic piece of work that you couldn't find one error on so they should get the A+; however, the other people who I'd given an A+ before there was a lot more errors there so I might blabla, what I would do is move them down. Or at least change the number, so if it was like 25 to 27 was an A+ / and I'd given them like 26, I might drop it back to 25.
> Helen: Yeah I think probably within the grade, yeah, I would've ranked, you know, the A's <...> / Yeah and I think it is important because if you don't have that sort of flexibility, then then you go 'Well no, I gave that an A+.' So, too bad ///. (Focus group B)

However, Helen also mentioned that she found "it was really difficult" to keep consistency.

4.1.3.2 Weighting grades

The interview data revealed that tutors tended to grade up when deciding which grade to award. For example, Mia believed she marked higher than other tutors because it boosted students' confidence:

> Researcher: Do you agree with the lecturers on the criteria?
> Mia: I think yeah, I always like the criteria, but I'm just not as heavy handed as some of the other markers were. I tend to mark a bit higher because usually if it is a first essay, it kind of

shatters confidence a little bit if they don't do really well. My students when they got their essays back they thought I'd marked them quite easily as well. They were very happy with what they got. (Mia, Interview)

George mentioned he marked higher for a hard working student. Similarly, Mark would weight grades according to students' starting points and efforts:

If it was an international student, you know, I will be a little bit lenient on them for the first two weeks but if it carries on I'm not reluctant to do it /// It depends on whether they put the effort in, you know, I will be nice to them. If they don't put the effort in, I'm not reluctant at all, I would give them a fail like that you know, and I have been advised to do that as well by the people above me like lecturers /// I wanted to be fair to everyone. (Mark, Interview)

Helen also believed the first assignment should be marked more leniently. She believed she was lenient in marking, but she also mentioned her experience of re-marking because she marked lower than the lecturer:

I probably read them through once, you know, all of them, and got some idea of how they were doing and how it was going. And then each of them then I would read and mark again. And then after I'd given some of them back to [[lecturer name]] and he said that I was being a little bit hard, so then I went through them all again to see whether I needed to increase their marks, whether based on what he'd told me whether I needed to increase their mark or whether I thought it was still consistent with what I'd done with the other ones. (Helen, Interview)

Three tutors mentioned they avoided giving fail grades: Jo was not sure whether to mark down or lower in case part of the answer was right. She would refer the assignment to the senior tutor rather than fail it herself; Mary claimed that she "refuse to fail people"; Frank had never given a fail grade. It seems that providing a low or fail grade caused emotional reactions among the tutors (c.f. Section 4.4.1.1). For example:

Researcher: Do you feel hard to fail anybody?
Henry: Yeah I – well, I have done it. I mean, I don't think anybody likes doing it. Um, and you sort of try not to do it, if you can. Sometimes you might just get a piece of work that is /// generally /// you know ///
Researcher: Yeah.
Henry: Awful. (Henry, Interview)

In the focus group discussions, tutors agreed that they tried not to fail a student. One reason for avoiding a fail grade was to encourage students to maintain their motivation:

> Mia: Th- the bare pass just to stop them being discouraged.
> Frank: Yeah the bare pass just to say 'Look I should have failed you but I'm gonna pass you, just work on your strong points.' I think maybe that'll < ... > in my mind that kind of might give them a bit of a kick along /// but, I don't know, me, I'm just too nice for my own good, who knows@. (Focus group A)

Later on, the financial consideration of a fail grade was raised:

> Mia: If you fail a course you have to pay for it and do it again. (Focus group A)

In Focus group B, the financial reason for not failing students was related to the commercial concerns of the department or university.

> Helen: Yeah, I think um I still probably – once she got here tended to avoid giving – yeah, I probably still wouldn't – I didn't, in fact, fail any of the second semester ones without talking to him. Yeah, I think um I wouldn't – I, I don't mind giving A pluses, at all, but I don't like failing people u-unless it's /// – um partly because you know I'm not sure, and partly because I /// kind of think the lecturer's the one who's going to wear it. [yeah] If someone complains, so they really need to take some / [have some input]
> Maria: [Actually have some] / some interesting comment from um – 'cause I talked to – 'cause I # didn't mark for our paper [mm] that we tutored on, last semester, and, but I talked to the lecturer about it, and she said she / tended to try NOT to fail people um for the reasons that she actually – I think she wrote something down here about um // uh / that they were very – that she was very aware of how that might affect their actual: [mm] – ah particularly their first year, their attitude towards university, their [mm] [yeah] . um social # effects.
> Helen: Mm // I think there's also commercial imperatives they want you to come back, don't they?
> Maria & George: Mm, Yeah. (Focus group B)

Besides, tutors also had different opinions on whether language should be marked down: Mark would mark down errors if they would appear repeatedly; George and Frank would not mark down for minor mistakes; Emma would not mark down for spelling mistakes; Maria and Frank would not mark down for language, whereas Henry mentioned he marked down unclear expressions,:

> I thought your ideas were good but you just – your writing was all, [mhm] your writing style was a bit messy, just, it was too difficult to sort of understand your ideas. Um, he- he'd been to

get a second opinion from another lecturer in the course? and then I'd marked it, and this other person has said 'Yeah, I think that seems ok.' But when I came to mark it maybe he thought // he was doing something or, you know, getting a better grade than what he actually got. [mm] Um I said to him, "Look if you want to go and have it remarked, by someone, obviously y- you can do that or that's fine, b- but he seemed to think that … I gave him a valid explanation of why my marking / – of why I marked the way I did, and I said 'It's your structure and your kind of sentences and that, I just, the expression was too difficult and I found it too hard to' – I could get that he had some interesting ideas? but I just found the writing difficult, and we had, you know, a grade sheet, so unfortunately it just turned out to be what I thought your mark was. (Henry, Interview)

4.1.3.3 Summary

This section reported tutors' beliefs of grading. The key findings were that tutors used different strategies to keep consistency in assessment and tutors tended to grade up and avoid giving a fail grade. However, tutors had different opinions on what should be marked down or up, which may be the potential cause of the inconsistency of assessment.

It was also found that criteria were used as a main tool to keep consistency of assessment and that tutors were concerned about students' emotional reactions to assessment. Both findings will be reported and discussed in the following sections.

4.1.4 Beliefs about Criteria

Concurring with the questionnaire respondents, all the interviewed tutors mentioned they used criteria or sample marked assignments provided for them by a lecturer or a senior tutor. Criteria were indeed the main tool to keep consistency of assessment among tutors. For example:

> Researcher: Do you refer to that guideline or criteria while marking?
> George: Absolutely.
> Researcher: You keep them at hand?
> George: Right there marking there.
> Researcher: So why do you keep them at hand or refer to them?
> George: I do that because of consistency. I want everything to be consistent. (George, Interview)

However, the criteria varied. For example, Cecile mentioned she used a clear structured marking guide, so both tutors and students knew the specific requirements. Helen mentioned the criteria were general and changed for different

types of assignments. Simon and Eva used sample marking. Emma had a marking schedule for one of the papers she tutored, which included right or wrong answers and points for each answer, while the criteria for the other two papers she tutored were about structure and language. Mark, George, and Jo also used marking guides. The tutors also held different views about the value of the criteria they used. For example, Anna followed the criteria even though sometimes she did not agree with them.

> Researcher: So you mean from the criteria, you can find the expectation of the lecturers towards the students. Do you think this expectation is the same as your expectation?
> Anna: My expectation doesn't count. Because in my initial work doing tutoring and marking I would take the work that I've marked back to the lecturer get them to mark it separately from me so that we could see that we were actually marking to the lecturers' expectations, it's not my expectations because it's not my course. I'm just there to support their goals. (Anna, Interview)

Tutors believed they had to use their personal knowledge to make a decision when assessing writing because fixed criteria could only serve as a point of reference. Mia mentioned the criteria she used were a bit 'linear'. She mentioned different papers put different values on expression. Besides the criteria, she had to use her own knowledge to make an assessment.

> Mia: The last ones I marked I did I have marking criteria. I was a little lenient with it but I used the criteria and then looked for what I would do in an essay because I know roughly what constitutes an A or B in [[department name]], quite a good judge compared to my own work, what I used to get. (Mia, Interview)

Frank used his intuition while marking:

> But, err, when you are marking you got to kind of, use your intuition you can't kind of follow instructions to a tee which it is like I don't know it's quite difficult but yeah I do follow the guidelines that my lecturers give me but I use my own intuition as well when I am marking, so, yeah. (Frank, Interview)

Jo also had to make her own decisions when the criteria could not advise exactly what mark should be given when evaluating ideas.

> Researcher: Do you use any criteria while marking?
> Jo: Yes, we do. We get given, what happens if [[name]] gives us a sheet that has the answers on. But the criteria is quite lenient and that [[subject name]] there is not always a right or wrong answer, it's quite qualitative and so it is sort of guideline, it

> gives us a guide as to figure out what the answers are, but like as I was saying before, it is very much I think it is right or wrong because it is often a wordy answer and I have to decide whether or not they have got the idea or not. (Jo, Interview)

George believed that this experiential knowledge or intuition was subjective, which therefore led to inconsistency:

> We have guidelines that say these are the things that you want to look at when you are marking. But it is up to us to decide how much or which ones we think more important than the others. And you see that's a subjective thing. (George, Interview)

In the focus group discussions, it seemed that criteria were not a major concern for tutors. However, the tutors used criteria as the standard which could justify the assessment when students came back to them and challenged the grade:

> You might get someone who challenges, and you actually should be able to defend in some respects your mark, you know, depend – usually there's like a marking criteria or something so it is really important. You can say, 'Well yes, they did that part right, but that part was like ///, you know,' so if somebody comes back to you, two come back to you with, you know, like, similar essays with different marks, you can say 'Well the reason for that is'. (Helen, Focus group B)

In sum, it was found that tutors believed that criteria were important for their work, a tool for them to keep consistency and justify the grades. However, they were aware that they needed to use their personal knowledge and intuition because criteria varied for different written tasks and criteria for written work were sometimes "linear" which were designed in a straightforward way thus might not be sufficiently helpful in addressing various issues emerged in students' written work.

4.1.5 Summary

This section reported key findings of tutors' beliefs about assessment and feedback. It was found that tutors believed good written work should address the topic and be clear. Moreover, content and clarity were also the two main foci of feedback. Tutors believed feedback aimed to help students improve but some tutors doubted the effectiveness of their feedback. Tutors believed positive feedback would be more effective and they tried to be positive by using praise, ticks, pointing out positive aspects of written work, and grading up. Tutors related both good written assignments and feedback to grading. Tutors believed

grading should be consistent. The main tool to keeping consistency was the criteria they were issued. Tutors also used various strategies to maintain consistency, such as making comparisons and reading several times before deciding on a grade. However, there was potential inconsistency in grading due to tutors' divergent beliefs of weighting grades. A common tendency of grading was that tutors tended to avoid giving fail grade because it may harm students' motivation or may have financial consequences for both students and departments. All these findings indicated that assessment rather than improvement was a major concern of these tutors despite their stated belief to the contrary. These beliefs will be compared with tutors' practices in the following section. Furthermore, tutors beliefs and practices revealed that emotion played an important role in their assessment activity, which point will be reported later.

4.2 Practices of assessment

Data relating to tutors' practice of assessment were collected by think-aloud and their reflections on practice during the follow-up stimulated recall sessions. Nine of the sixteen interview participants (Helen, Anna, Mia, Simon, George, Mark, Jo, Cecile, and Emma), voluntarily participated in think-aloud and stimulated recall sessions.

The participants' actions during the think-aloud session were reading, thinking, marking, and writing feedback. Their verbalization included reading aloud the assignment, repeating some words, commenting on the content and writing, recalling the students' general performance, relating the assignment to the criteria, recalling the content and focus of the lecture and tutorial which the assignment was based on, repeating the oral instructions of the lecturers or senior supervisors, identifying the errors or good aspects of the assignment, and reading aloud their feedback. Eight of the nine participants started to read and mark from the beginning of the assignment; one participant looked at the reference list first. Some read aloud most of the assignment while marking, while others read silently. The majority of participants referred to marking guidelines or sample answers; two tutors did not bring the marking guidelines with them as they said they had the criteria clearly in mind. Eight participants used pencil and rubber whereas one participant used red pen. All participants thought silently before they verbalised their feedback. Some error corrections like the spelling mistakes and the routine

feedback such as "Good", "Well done" were spoken aloud simultaneously to the participant writing. However, there were relatively longer periods of silence when a participant was thinking about the sentence structure or giving longer feedback; especially as regards negative aspects. Most participants spoke in a low voice while they were marking.

This section will report the key findings of practices in comparison with the beliefs reported in previous section. The structure of this section is Section 4.2.1 Good written work; Section 4.2.2 Feedback; Section 4.2.3 Grading; Section 4.2.4 Use of criteria; Section 4.2.5 Summary.

4.2.1 Good written work

More convergences than divergences were found between beliefs and practices of the nine tutors regarding what constituted a piece of good written work.

Table 2: A list of assignments marked by the tutors during Think-aloud sessions

Type of assignments	Tutor	Number of written work
Lab reports (short answers and descriptions)	Cecile	2
	George	2
	Mark	2
Review	Anna	2
	Simon	1
	Emma	2
Short answers	Jo	2
Essay	Helen	2
	Mia	1

Convergences between beliefs and practices were found in three aspects: firstly, all the tutors who participated in the think-aloud and stimulated recall sessions assessed the content of the written work.

> 'Good'. So she is looking at both situations. (Mark, Think-aloud)
> I like that idea. I'll give it a tick. (George, Think-aloud)
>
> If it's a general thing, like she's being very general with some of
> her ideas, I'll write that in the comment box, saying 'you need to
> be more specific and expand your ideas' (Mia, Stimulated recall)

Secondly, six interviewed tutors believed the language should be clear (Helen, Simon, Mia, Mark, George, Emma). This issue was mentioned by six participants in their think-aloud and stimulated recall sessions (Simon, Mia, Mark, George, Cecile, Emma) (c.f. Section 4.2.2.2), for example:

> Researcher: When you feel the sentence is not clear, can you understand them?
> Emma: Sometimes I have to struggle. Read the sentence three times and found I corrected it in the wrong way so I have to give up my correction. (Emma, Stimulated recall)

It seemed that errors at lexical and syntax level attracted the tutors attention in the process of marking because the errors were the main sources of unclear expression for any type of written assignments which influenced the understanding of content.

Divergences between beliefs and practices of good writing were found in two aspects: firstly, feedback on punctuation (such as the use or misuse of commas, apostrophes, and quotation marks) was given by four participants (Helen, Anna, Mia, Emma). This practice was not mentioned by these tutors in the interviews. Secondly, paragraph structure was mentioned by Emma in the interview. In practice, this aspect of writing was addressed not only be Emma but also by three other participants while were marking reviews. Focus was on the paragraphs of introduction and conclusion of assignments, for example:

> Researcher: I suppose you spent more time at the beginning of this assignment. I remember that you …
> Emma: Yes. The introduction and the conclusion usually take the longest time to read. And they usually should take the longest to write, because you're trying to include a lot of information in them. You're kind've trying to open your assignment and say this is what I'm going to be talking about. And it's in that part that you're trying to establish what your argument is and also to draw the reader in. (Emma, Stimulated recall)

It seemed that the standard of good written work was decided by the items listed in the criteria:

> So going through it I can see that he's actually started to have quite a a good introduction, which is a requirement, um, //// but um, already I'm seeing that he isn't ah / putting in correct um / citations /// [[speaking while she writes]] (Anna, Think-aloud)

> Ok so, in the assessment sheet /// they would say that the first part of it is /// done very well //// And as is the second /// This part's done very well, the analysis. And there are very few grammatical errors. . So, I would say to give a general comment, I would say / 'It's is an excellent / essay /// with / great / use of / scholarly / sources to support / a coherent / argument". [[speaking while she writes]] (Helen, Think-aloud)

4.2.2 Feedback

In comparison with the interview findings, the key aspects of the practice of giving feedback are to be presented in the following sequence: the purpose of feedback, foci of feedback, positive and negative feedback, and summary.

4.2.2.1 Purpose of feedback

In the interviews, the tutors believed the purpose of feedback was to help students improve. It was found in their practice that error identification and correction was used by tutors in the first instance to enable them to understand the content. For example:

> Again, I- all the way through I'm having to / rewrite sentences // to / be able to interpret them /// though. (Anna, Think-aloud)

> So I just go through // and / pretty much say the sentences how I would say them or kind of try and get some understanding of what they're trying to say, and at times it's easier than other times. There are times you can just move a few words around to get a sentence, but other times you really do wonder what the point of that paragraph was. (Simon, Think-aloud)

Later on, in her stimulated recall session, Anna explained that error correction was intended to help students improve their grammar:

> So that's the sort of grammatical work that I am doing, (yeah) framing it, helping them to think about how you frame a sentence better or. (Anna, Stimulated recall)

However, the main purpose of giving feedback seemed, to most of the participants, to be to justify a grade:

> I mean I don't always tick what they do, it depends, yeah I mean, but I always – but if it's one section like this, I'll always circle or cross what they haven't done, so like, (yeah) you know, so so they know where they lost marks. (Cecile, Stimulated recall)

This point was made even clearer by Simon:

103

Yeah. I mean at times if / if I was extensive, here and here [[pointing at in-text feedback]], I would just write, 'See essay', and leave that, but only this is more like a justification of the grade, and it's similar – this is a similar format for everybody. (Simon, Stimulated recall)

Moreover, it seemed that the amount of corrective feedback was constrained by time available for assessment:

But it is not really my job to teach them how to spell, and put full stops in sentences or commas in the right places. So I sort of just leave that otherwise it would take me a lot longer to mark. (Jo, Stimulated recall)

Although no suggestions on how to improve was observed in the think-aloud sessions, some of the tutors afterwards talked about the improvement they observed in students' assignments. For example:

Researcher: Have you noticed they have made any change?
Helen: No they're getting worse! @ Yeah, nah actually the changing probably would be err, / in / the referencing? Some of them who at the beginning weren't reference now reference. So that's really good because that is important for doing [[subject name]] through, you know, you have to reference in [[subject name]]. And yeah, so that's really important. So that's good. That some of them who in their first assignment did no referencing now know how to do it. So that's really good. (Helen, Stimulated recall)

Simon mentioned the improvement of transitional words in the assignment he marked:

…like here I like how they've actually got a good linking phrase, 'in addition', you know which kind of works, but you know in my previous marking I've never really commented on that stuff. (Simon, Think-aloud)

It seemed that improvement could be observed in some assignments but not in others. Some aspects of writing, such as reference or transitional words, were noticed by participants as improvement, which may be the effect of feedback. This confirmed the point made in the survey, where 40 per cent of the respondents felt that students did not read their feedback carefully. Further confirmation was found in interviews (c.f. Section 4.1.2.1) and in one of the focus group discussion:

You've been here for a whole semester and more you still haven't got it. (Helen, Focus group B)

4.2.2.2 Foci of feedback

Convergences between beliefs and practices were found in two aspects: grammar and vocabulary, and positive and negative comments.

Firstly, all participants except two (George and Jo) said they gave feedback on grammar and did so in practice. Five participants (Helen, Anna, Simon, Mia, Emma) gave feedback on almost all the errors they came across while marking. As was expressed in the interviews, the aspects of writing participants commented on in practice were mainly at lexical and syntax level. One tutor, Simon, spent the longest time (53 minutes) among the nine tutors in assessing a three-page-written work by a non-native English speaker student, forty-two minutes of which was spent in reading and correcting the grammatical errors:

> Ok. So this is like what really slowed me down doing it. I have to go through and do it like that. I mean I know I don't have to, and I do it far more than what anybody else does. But er, it is important to me, as a part of the marking process at the moment, I can't really do it intuitively, I have to kind of think through slowly like that. (Simon, Think-aloud)

Secondly, all except one tutor (Jo, who assessed short answers to questions) commented on both positive and negative aspects of the written work in their in-text feedback. Six participants also commented on both positive and negative aspects of the written work in the overall comments. For example:

> So I think that she's sort of in around seventy / four, which is / um a B+. [[pause, 13 seconds]], quality of writing, [[pause, 16 seconds]] the quality of writing went down. [[Reading the overall comment]] '[[student name]] Your review gave a very good summary of the movie, you understood the core message of the film and draw on the effective sequences to demonstrate the brutality of genocide but [[further comments]] You need to reread your essay, some of your sentences are grammatical incorrect. Well done. [[tutor name]]'. (Anna, Think-aloud)

This point was confirmed in the stimulated recall session:

> So you are trying to reinforce what they have done and then show – demonstrate what they haven't done / and um, / and because it's the first year, first essay, you try to be quite positive. [yeah] To affirm what they are doing? Rather than going 'nah! that was absolutely terrible'. (Anna, Stimulated recall)

One divergence between beliefs and practice was that although Mark, George, and Cecile believed that when they were undergraduates themselves, they benefited from plenty of feedback given on their assignments, they did not give a great deal of feedback on the reports they were marking. For example, the overall comment Mark gave on the first assignment was:

> I'll say er 'Great effort' [[turn to the first page for the name and then write it down]] [[student name]] Err, 'Try and use more ///// of the social science terms and concepts.' That's one done there. (Mark, Think-aloud)

The overall comment on the second assignment was:

> I'll say 'Good effort,' err, [[reviewing the work]] 'Good effort,' umm, 'probably more use of /// the social science terms and concepts /// and um try [[thinking and writing]] to answer the questions in complete sentences. This will help you /// BIG TIME for the essay and exam.' (Mark, Think-aloud)

It seemed that tutors' practices regarding overall comments were constrained by the tools they used. For example, Mia did not provide an overall comment on the essay she marked because she used an assessment sheet which divided comments into sections. The type of assignment and the criteria that the tutors used may also have limited the amount as well as the content of the tutors' feedback. The feedback of those who assessed lab reports consisted mainly of marks, symbols (e.g. ticks), and brief comments such as encouraging words on the each item listed in the analytical marking schedules. George did not provide an overall comment to one lab report he marked. Similarly, Jo did not provide an overall comment for the short answer questions.

4.2.2.3 Positive /negative feedback

All the tutors provided positive feedback in practice, which was convergent with their beliefs. The techniques of being positive mentioned in interviews (ticks, praise or encouraging words, positive aspects, grading up) were all found in the think-aloud data. All the tutors used ticks when providing in-text comments. Besides, some tutors (Cecile and Mark) also used smiley face icons in their comment as a way to be positive. No harsh expressions were observed in tutors' feedback. For example:

> They haven't put in a full answer but I still give them a tick for it, because they just about got it right. (Jo, Think-aloud)

That is a good argument too. A very good argument. So I'll tick that twice. (Helen, Think-aloud)

I usually just put a tick next to it – where the bulk of it is, or at the end of the paragraph. Er, and if it's an especially good point I 'll put an explanation of why I put that tick there. (Mia, Think-aloud)

I'm saying 'Well, good work [[student name]], /you've covered this section well.' [[pause]]   marking guide, // for um, clarification on any lost mark, // any marks lost /// 'Good job' and a smiley face. (Cecile, Think-aloud)

In the overall comments in stimulated recall sessions, tutors further explained their beliefs about and practices of being positive. For example:

Because Jan tells us, it's better to be positive than negative. Like don't mark – don't mark their work negatively because // it'll give like, sort of bad impressions of the department?, like we're really negative and we want to show people they're wrong – we want to show them that they're almost right or they're right. (Jo, Stimulated recall)

Well, that that was a very good essay. Good essays are easy to mark. They're /// they're – um you read them and you think, 'Yes! They got it.' You know. Yeah, um and I find it easier to write positive comments than to write negative comments. I I find writing negative comments – I don't want to be too negative. (Helen, Stimulated recall)

If it's a general thing, like she's being very general with some of her ideas, I'll write that in the comment box, saying 'You need to be more specific and expand your ideas' and stuff like that., but I'll also give like positive things like 'I really like what you've done with this section', and um, I'll suggest things that will make it stronger. (Mia, Stimulated recall)

Um, I usually just like to put it just so there'll be some sort of positive thing in there, you know, regardless of whether they've done a good or bad assignment I usually put a smiley face – it's kind of just a little signature thing I do, I think. (Cecile, Stimulated recall)

4.2.2.4 Summary

Tutors' practice of giving feedback was reported in this section in comparison with their beliefs. It was found that the major convergences included the following aspects: the good written work should be clear and address the topic;

tutors focused on content but all gave feedback on grammatical issues; and tutors tried to provide positive feedback. The major divergence between beliefs and practices was the purpose of giving feedback. Another divergence was the length and content of feedback: some tutors provided an overall comment while others did not; Some tutors covered both positive and negative aspects while others only used encouraging words. The divergences between tutors, and between beliefs and practice were related to the type of written work, the type of criteria and assessment sheet.

4.2.3 Grading

As was found in the interviews, most (six out of nine) participants provided both feedback and a grade when assessing students' written work, except that Jo was told by the senior tutor not to write down the grade on the assignment, and Emma would give a grade later on the basis of comparing a group of assignments.

4.2.3.1 Consistency

One convergence between tutors' beliefs and practices was that they tried to keep their assessment consistent by comparison or several rounds of reading. To provide a fair grade, six participants mentioned they compared the marks of several assignments. For example:

> Researcher: Oh yeah. Is it a general practice to read the assignment twice?
> Mia: Err, Yeah usually, yeah. Usually when I first start marking, I'll read the first essay I mark at the end just to make sure. Um, I usually do a group of ten or so, and I'll do – I'll look at the first one quickly, and then I'll do the rest of them in a group and then the first one again.
> Researcher: And then for the second round of marking, you
> Mia: I'll just do the same – do that same kind of thing again. Um, so just to make sure I don't mark the first one too harshly or too easily depending on the others that I marked. (Mia, Stimulated recall)

Emma mentioned that she marked in the same style as the lecturer.

> In this particular one, it's influenced by the person who / marks the other assignments, by the lecturer – because I've had her for a lot of papers, so I just mark in a very similar style to how she marks. [ahh] It's so there's level of consistency… (Emma, Stimulated recall)

However, Emma did not award a grade because she usually read through the written work and provided in-text feedback for the first round of marking, and then decided on grade and overall comment. A similar approach was used by Anna but she had adjusted her way of marking over time (c.f. Section 4.2.3.2).

However, it seemed that inconsistency was unavoidable: Simon mentioned he might mark higher or lower than the lecturer:

> Simon: I'm just trying to find the lecturer's marks as well because he attends the tutorials also.
> Researcher: Is it a sample?
> Simon: Err, just his, his marks on those presentations and I am allowed to agree or disagree. I use to mark a little bit higher I think. I mean higher and lower. (Simon, Think-aloud session, before started to assess the assignment)

Mia mentioned the efforts taken to keep consistency by marking together with other tutors of the same course:

> We just pick out a best one, a worst one and something in the middle, and we'll read them and see whether we're on the same page, and if we're on the same page, fine, we won't have to re-mark anything, um sometimes they'll get pushed up or pushed down because we have different ideas, but otherwise it's – we're both this even now. (Mia, Stimulated recall)

Both Mia and Jo mentioned they might not be able to be consistent regarding the amount of feedback they gave to the first and last assignment because of time limits or because they might be tired towards the end of a period of marking. For example:

> I think kind of at the start when I'm marking, like, you haven't been doing it for very long, and you can – you have the effort, and you can be bothered to quickly write something in, but maybe on the last one, I wouldn't do that. Which is kind of unfair for the students, but it's just kind of what happens I think. [Yeah] Yeah, you have the extra effort at the start because you've only just started doing it, but at the end, you're kind of thinking 'Oh, I've been doing this for hours, I just want to get it finished. I'm not going to hold myself up anymore by writing.'
> (Jo, Stimulated recall)

Emma put it this way:

> It is up to you to decide sort of how well they have managed to answer it. (Emma, Stimulated recall)

4.2.3.2 Weighting grades

Convergences were found between beliefs and practice regarding weighting grades. Firstly, participants hesitated to make a decision on what grade should be given when there were choices of whether or not to fail an assignment, or when the mark fell between two grades. No fail grade was observed in the think-aloud sessions. For example, Helen did not write down a mark on the assignment that she thought might be failed:

> As far as marks go, I would want to give it a good mark based on its argument but in terms of /// the amount of referencing that it hasn't done, it would certainly lose a large amount of points so /// my thinking would be /// perhaps a B / Yeah. (Helen, Think-aloud)

> Ok, so do you need to know that / um, after I've marked these, I'll take some to [[lecturer name]] to confirm that our marks are consistent. This would definitely be one I would be taking to [[lecturer name]] because / my instinct is to fail it because of the complete lack of referencing, but it is a very good essay. (Helen, Stimulated recall)

Mia also had the same practice:

> So if there's anything that I get and it's a fail, I won't put any grades on it, I'll just put my comments. And then I'll let the lecturer decide what mark to give them. (Mia, Stimulated recall)

Secondly, three participants (Simon, Helen, and Jo) showed a tendency to mark higher when there was a possibility to mark lower. For example:

> I mean I probably/ yeah, / I'd probably want to give it a B, but I will talk to him [[the lecturer]] whether – how much weight he is going to put on the lack of referencing really. It's entirely possible that some lecturers would fail it, but, but yeah I would need to talk to him about it so that we're consistent across the course. (Helen, Stimulated recall)

> I had to look at it more closely, which is actually quite frustrating when you're marking because it holds you back / a bit, or for example when a student has really messy answers? it is quite annoying, but I don't mark them down for it or anything. (Jo, Stimulated recall)

4.2.4 Use of criteria

Convergences between beliefs about and practice of using criteria were found in two aspects: the application of criteria or sample marking sheets and the variety of marking schedules.

First, criteria or sample marking sheets were used by all participants except Simon and Emma. For example:

> Now I'm just going to check the marking guide and see if there was anything she could've mentioned or that she's missed out. (Cecile, Think-aloud)

Secondly, criteria or marking schedules varied in style, content, marks given to expression, and the space for feedback. The five tutors who assessed essays or reviews used holistic criteria; the four tutors who assessed lab reports or short answers used analytical marking schedules. It seemed that tutors interact more with analytical criteria than with holistic criteria while reading and grading:

> So I'll just look at what he's got, / um, he hasn't given that information, / in particular that // quite a bit of that that section, I'm going to give him 1.5 out of 6, he's missing quite a lot of it? / er, just checking for grammar and spelling ////I'm going to take off a quarter for the grammar – he's made a couple of small mistakes. He didn't keep it all in past tense so I'll take a quarter off (###) half a mark of that anyway. So I'll just add it up out of 14 //// and 10.25 out of 20 / and I've just write 'Good' and a smiley face. (Cecile, Think-aloud)

> Er, they got two wrong there, so I will give them five out of it six // (#######) [[reading]] Good. I'll give them three point five, um / , OK, (##). Yes. Yes. Yes. Um, No, that's wrong. Er, it is (#). Yes. Yes. No. No. Yes. Yes. No. No. Not [[reading]] (##). Yes. Yes. Yes. Yes. Good. (#####) There we go. There's the key word 'non-renewable'. [[reading]] (####) Just put down here. Yeah. Yeah. Yeah. Yeah. OK. So no. Yes. Yes, both, both yeah, and this, both, so that. Yeah. Yes. Yes. Yes. Yeah. Yeah. Good. Ok now, and asking them to be reflective, yeah, mm [[chuckles]] that is good. Yeah. Good. Good. Good. Right now, get it done. Why use my brain when there's a calculator [[calculating]] So this person got fifteen point five / for that. (George, Think-aloud)

Holistic criteria were often used after tutors had read through the written work:

> I'd probably say this one's probably one of the better essays I have read so far, and they, and they followed the criteria that's been laid out to them. And just looking at the essay and looking

> at the marking schedule, I'd probably easily give them a B. (Mia, Think-aloud)

It might take cognitive effort for weighting grades:

> I'll put it in the C range. Maybe in the // in the, you know, you have to # break some of it, but / um / may be around, around 58 / C // breaking it down / quality of writing out of 20 //// be / half way there // organization … (Anna, Think-aloud)

In the stimulated recall session, Helen mentioned the longer the assignment, the more explicit the marking schedule. Cecile mentioned the marking schedules for the second or third year papers were not as specific as those for first year papers.

Divergences between beliefs and practice were found in two aspects: not carrying a marking schedule and not following the schedule closely.

First, Simon and Emma did not carry the marking sample with them while they were grading during the think-aloud session. For example:

> Researcher: You do not need criteria.
> Emma: Umm no. She provides us with examples of ones that she's marked. And then – to give us an idea of kind of the / level of comments that she wants us to make. [ahh] And then I went and checked with her yesterday that I was doing the right thing.
> Researcher: So you are not given a written document thing?
> Emma: No, because / it doesn't really work for some of them. So ///
> Researcher: Didn't really work for / for?
> Emma: Well, it would depend what the specific criteria is, because I mean, it's //// I don't know, it's difficult to tailor criteria to the /// [[silence]].   (Emma, Stimulated recall)

Secondly, participants may or may not completely follow the criteria regarding formal issues. Language was not an item that was listed in the sample marking used by Mark, George, and Jo, who marked lab reports. However, Mark gave feedback on incomplete sentences, while George and Jo did not give feedback on writing. Grammar counted for 1.5 marks in the criteria used by Cecile and she pointed out grammatical problems such as the use of wrong tense in one assignment. The five participants (Helen, Anna, Mia, Simon, and Emma) who marked essays or reviews all paid much attention to grammatical errors, even though it was not a main aspect of marking according to the criteria. For example,

in stimulated recall session, Anna and Mia all recalled they corrected almost all errors in the assignments they marked. For example:

> Researcher: Yeah. It seems you read every word and you correct every error?
> Anna: Err, not always. I mean, well, Yeah. (Anna, Stimulated recall)

> Researcher: Actually you tend to correct all the problems.
> Mia: Yeah, um usually I'll go through and correct as many as I can. Sometimes I get a little lazy and if I've done it on every page of the assignment I'll tend to stop after a while, [yeah] because they'll get the point in the first couple pages what they've been doing wrong anyway. (Mia, Stimulated recall)

4.2.5 Summary

This section reported tutors' practice of assessment and feedback in comparison with their beliefs. More convergences than divergences were found between beliefs and practice. The tutors used criteria but also have to use their personal knowledge or instinct in assessment. They assessed content of the written work and used similar techniques to be positive and consistent when giving feedback and grading. They tried to avoid awarding a fail grade. However, they had to address grammatical issues in order to assess content.

The major divergence was the stated purpose of feedback for improvement and the actual purpose served by their practice to justify a grade. It was found that contextual factors, such as the type of written work and criteria, the quality of the written work, and tutors' physical conditions and beliefs, interacted with each other and influenced tutors' practice of assessment and feedback.

4.3 Sources of beliefs and practices

The sources of tutors' beliefs of assessment and written feedback were mainly explored in the interviews. Moreover, the data relating to the sources of beliefs were further triangulated by think-aloud and stimulated recall sessions and focus group discussions. It was found that tutors' beliefs were derived from their participation in the activity of assessment and other teaching and learning activities that were relevant to assessment across time. These sources of beliefs will be reported in this section in the following sequence: 4.3.1 Knowledge distributed in the community of practice; 4.3.2 Interactions within the community

of practice; 4.3.3 Reflections on personal experiences of assessment; 4.3.4 Contact with members of other communities of practices; 4.3.5 Summary.

4.3.1 Knowledge distributed in the community of practice

The formal requirements of assessment was the knowledge distributed in the community of practice. Tutors were required to provide written feedback when assessing students' written work. The written feedback usually included both in-text feedback and an overall comment, although some tutors provided only an overall comment while others gave in-text feedback, with or without giving oral feedback. The following requirements for giving feedback were found in the interview data: the work of assessment and payment, reviewing drafts, error correction, overall comment, grading, and criteria.

4.3.1.1 Work of assessment and payment

As noted in Chapter One, the part-time tutors worked in a hierarchically structured community of practice. Generally speaking, undergraduates' written work was assessed by lecturers or professors who taught the course. In cases where there was large enrolment in a course, part-time tutors were temporarily employed to assist lecturers or professors to provide tutorials and assess students' written work. These tutors were supervised by the lecturers or senior tutors. In addition, tutors' work was subjected to formal student appraisals at the end of each course.

The interview data revealed that the responsibilities of the part-time tutors varied. One part-time tutor (Anna), was only employed to mark assignments, while other part-time tutors also provided tutorials to students; two tutors (Mia and Emma) were employed to assess students' written work from different courses across departments; another tutor (Cecile) had more responsibilities than tutoring, such as administration and lecturing.

The assignments that the tutors assessed were written by students including both native and non-native speakers of English who had a wide range of academic and literacy backgrounds. Most of the survey participants (75%) agreed that native speaker students had fewer problems in writing than non-native speaker students. This was further confirmed by interview data. For example:

…I knew that person and I knew that he'd tried very hard and he was also an international student. And I know that he really struggles with the language and things like that. And there was a lot of language in this article that would have been confusing and ambiguous. So I kind of took a little bit pity, and I probably shouldn't, but the thing is you know you want to encourage them to do well… (George, Interview)

For the same reason, Simon spent the longest time (53 minutes) of the nine think-aloud participants in assessing the written work of a non-native English speaking student. In his stimulated recall session, Simon confirmed it took him a longer time marking written work of non-native speaker students:

Researcher: Is there any difficulty if you mark an assignment written by a second language learner?
Simon: Em, yeah. I should say it takes me much longer. (Simon, Stimulated recall).

The students whose work was being assessed by these tutors had different academic backgrounds. This point was raised by Helen:

Some of them have come with a lot of [[course name]] at high school, some of them are law students who are doing one paper in [[course name]] and are doing it just doing it just for fun. There are some from teachers' college doing it. So yeah, So it's sort of, you have a wide range of students with a wide range of sort of backgrounds. (Helen, Interview)

There was consensus that their students had difficulty in writing effectively. For example, the written work of first year undergraduates had many problems because "school doesn't prepare you for everything" (George, interview). This point was also supported by focus group data:

Henry: …Writing style is pretty poor / for most students, particularly in like first year …
Mia: they're, they're, they're not teaching grammar and syntax in schools anymore.
Frank and Henry: Yeah. (Focus group A)

The payment for part-time tutors was $16 per hour. The time paid for marking a piece of written work varied: 30 minutes according to Jan and Anna, 20 minutes. For example:

We've got a policy of giving the markers half an hour to mark an essay. I think that's probably quite generous. (Jan, Interview)

They allocate 20 minutes per assignment. I will take more than that. (Helen, Focus group B).

Like Helen and Anna, all tutors in interviews believed they spent longer than the paid time. For example:

> Researcher: How long will it take you to mark each assignment?
> Anna: I think the pay rate is that you get half an hour per assignment // so that will be forty hours per assignment. But I mean they often take longer than that. Depending on what it is, whether or not you have to rewrite them to, and expand them…
> (Anna, Interview)

Data from a focus group discussion also supported this point:

> Henry: Is the next one about timing, /// um being an old coot, can think back to th- a bugbear among heaps of tutors / about doing // – yeah you know you only get paid for a certain amount of time but obviously it always takes you – most of the time, a lot longer / [mhm] // to do all marking duties [mm] // depending on the personality of the tutor, some people could be fine // annoyed about it and sort of complain . .
> Frank: Yeah, I, I tend to find um, // ah when I first started tutoring anyway, # they give you like an hour for preparation and marking but I found I'd spend about / I suppose three hours preparing for a class, yeah, and now I'm just a little bit – maybe ### 'Ah who cares, I'll just go in there and do something'.
> (Focus group A)

However, it seemed that the time spent on marking varied, depending on the type and quality of written work, criteria and individual styles and requirements of assessment. In think-aloud sessions, the nine participants spent different times on marking a student's assignment, ranging from 9.5 minutes (Jo, lab report) to 50 minutes (Simon, essay).

4.3.1.2 Reviewing drafts

According to the survey, more than 60 per cent of tutors gave feedback on drafts. However, other sources of data revealed that feedback on drafts was not usually given by tutors. In interviews, six participants mentioned they did not give feedback on drafts (Helen, Anna, Simon, Frank, Emma, Martin). For example:

> Researcher: Do they submit drafts?
> Emma: No. Not for this one they don't. (Emma, Interview)

The reason for not accepting drafts was mainly because of local assessment regulations or conventions:

116

…for the written assignments I've got, the department had a rule about that we're not allowed to look at er draft work er we are allowed to discuss the students' ideas with them. (Frank, Interview)

…because we are not paid to do that. Our supervisors are paid to do that and they are quite happy for students to go to them for that service. (Martin, Interview)

However, the rules may change according to different assignments. For example:

Helen: Two students actually have asked me whether they can submit draft assignments. We decided not for this first one because I wanted to see where they were to start with without correct // I am not sure. I need to talk to [[lecturer name]] about what he thinks about that. Perhaps for their final assignment, which is a big essay, it might be a good idea just to make sure that they're getting their ideas that they need to get in order ///
Researcher: So you mean before the final submission of the final assignment they may submit a draft.
Helen: Yeah. But I need to talk to [[lecturer name]] about that, he'll probably tell me it's a lot of work. But, you know, I don't actually mind that. As long as it's in keeping with his policies. You know, because it needs to be fair for all students. (Helen, Interview)

Moreover, some tutors could decide for themselves whether or not to receive drafts from individual students. For example:

We weren't allowed to accept drafts last time, but I have mentioned in my class this year that if they need some help, because the reading report is quite short, I would look over drafts and just point out where they needed a bit more help…
For larger essays, anything over about a 1000 words I don't want to accept drafts. Because if the whole class does it, that is a lot of reading that I have to do.   (Mia, Interview)

Only Cecile said she gave feedback on drafts:

Researcher: Do you give feedback on drafts?
Cecile: Yes. We do have students bring to you. Again I write on them, and talk to them because they usually do what I mean. (Cecile, Interview)

It seems that whether or not to review drafts depended on the time that tutors could afford:

Researcher: But you don't receive any drafts from them?
Henry: Er, that – used to. That that's kind of changed a bit I think for this this paper because there's so many students involved. (mm) Um, and often the tutors are all doing other

117

work?, so the lecturers just said, 'Look it's too-, it'll be too busy' – you might have tons of students giving you all these drafts, and they expect you to mark them, and then often at – particularly at first year level, you can end up sort of writing the essay / for them if you like?.

In addition, Simon believed students would not usually have time to submit drafts:

People wouldn't take the soft deadline of drafts seriously, and they wouldn't have the motivation to have something completed by that time. (Simon, Interview)

It seemed that this opinion was also supported by Eva:

Researcher: Did students all submit their drafts before this assignment?
Eva: Three people. Not very many. (Eva, Interview)

In think-aloud sessions, only Mia mentioned that she had reviewed the draft of the written work she was marking. It seems that tutors were not required to give feedback on drafts.

4.3.1.3 Error correction

According to the survey, the most common type of feedback was underlining errors and providing corrections (71.43%).

According to the interview data, all tutors provided feedback on grammatical errors, although some tutors were not required to correct errors, the reason for which was explained by a senior tutor:

I do not encourage my tutors to correct every error because we cannot afford it. And the students wouldn't look at it anyway /// So we don't correct things. One of the reasons for that is that my tutors are only twenty one years old. And they are not trained teachers. // So they haven't done teacher training. Neither have they got particularly high English kind of qualification. They would have done English at school, but they haven't done English. So they are probably not a lot better than a lot of the students in class, so the lecturers wouldn't be happy if I was asking those tutors to go through and correct everything. They are not up to the sort of standard of [[names]] for example /// My tutors don't. So my tutors know if it looks good, reads well, and they can see some errors and they do, I encourage them to correct things but not literally word for word. (Jan, Interview)

Two tutors mentioned they were required to correct some of the errors:

Researcher: If there are some grammatical mistakes, are you going to correct them or?

> Eva: Again, what we have been advised from the lecturer, is to do the whole piece is going to take a long time. So she suggested what we do is that we do the first couple of paragraphs, and then write a list of comment errors like apostrophes, and different constructions whatever, /// at the end ///. (Eva, Interview)

> Researcher: If you find some grammatical mistakes, will you correct them?
> Anna: Yeah. Usually one of the lecturers I worked for he recommends if it is really bad, just do it for the first page and then just make a comment, this is something that you need to address, but at times I have to rewrite whole paragraphs just to actually make any sense out of just to be able to mark it. because you cannot get a sense you know they are trying to tell you something, but the way they've written it has make it so hard so you might actually rewrite this whole paragraph and say this is how you should have written it or could have written it. (Anna, Interview)

Other tutors did not mention any specific requirement for error correction. It seemed that the rules for error correction were flexible. It was usually up to tutors themselves to decide on whether to correct errors or how many errors to correct. In short, they corrected errors according to their own beliefs and the time available:

> Researcher: If they use a wrong word, would you correct it?
> Jo: Sometimes I cross it out, and I write what they should have written because I want them to know they were wrong and they should have written this because that's how they learn. But not all the time because that would be very time consuming @. So I just, you know only do when it is really a bad mistake, or really obvious mistakes. (Jo, Interview)

Four tutors (Jan, Anna, Mia, Martin) mentioned that they could recommend students to go to Student Learning Support Centre in the university if their grammar was really bad.

4.3.1.4 Overall comment

According to the interview data, the convention for the overall comment was to be positive.

> I will have written that on the document I am going to give you. I think it says when you are marking the essay, be very polite and be modest, and say positive things. See if you can look for the best things in the essay and say what is good, and then not what is bad or wrong or negative, but what you could do to make it better. I think that is quite fundamental. (Jan, Interview)

There were no other specific requirements for overall comment for most tutors. However, two tutors mentioned some requirements of the overall comment during the Stimulated recall sessions:

> …and [[lecturer name]] expects about that much comment. He doesn't want just a / a couple of lines. He's expecting a good / long comment. [yeah] So you know, so there is good feedback?, [ah yeah] so you know – I remember one of the ones he gave back to me and only had about four lines, he goes 'Well, this is like the minimum that you should do,' so he is wanting a more in-depth ///. (Anna, Stimulated recall)

Another tutor, Helen, mentioned that the lecturer reminded her of the word choice in the overall comment.

4.3.1.5 Grading

The tutors felt that it was important for their grading to be consistent. They might have to re-mark the written work or adjust the grades to maintain consistency of marking within the same cohort of students and with that of the lecturer:

> Researcher: So you mean you read a sample of assignments, several assignments, and then get an idea of their level.
> Helen: Yeah. So I read through again the ones that [[lecturer name]] gave me back, that he said he thought was a little bit hard, and then sort of read some others that I thought were at the similar levels to see whether I thought that he would have thought they were marked hard too. I found it quite a complicated thing. (Helen, Interview)

A similar experience of re-marking was mentioned by Mia in her stimulated recall session:

> When I was marking my first assignment, they didn't tell me what they wanted. So er, my marking was off compared to the other ones, so we had to re-mark everything. (Mia, Stimulated recall)

4.3.1.6 Criteria

According to the survey data (Table 3), most tutors received criteria, written and group verbal instructions from lecturers (over 89%). Most tutors used two or more forms of criteria, perhaps different criteria for different type of written work.

Table 3: Tools from lecturers on how to mark students' written assignments

Category	Number	Written instruction	Check List	Listed Criteria	Verbal Instruction		Sample Feedback	Model Answer	Other
					Individual	Group			
Total	28	26	16	25	17	25	13	15	0
Total	%	92.86%	57.14%	89.29%	60.72%	89.29%	46.43%	53.57%	0

The interview data confirmed that the most common types of assessment tools were marking schedules and criteria. The interview data also demonstrated that in some departments, tutors had group meetings before they marked assignments, during which the senior tutor would explain the requirements of assessing specific assignments.

Data collected by think-aloud sessions revealed that tutors who marked essays usually used marking schedules which included general items of assessment such as content, organization, argument reference, and expression.

However, it was also found in think-aloud and stimulated recall sessions that different tools were used for the same type of assignment in different courses. The criteria tutors used to assess reviews varied in descriptions of the requirements for each item and the marks allocated to each item also varied in different criteria. When assessing lab reports, samples of marking were used by tutors in one department, while analytical criteria were used by a tutor in another department.

4.3.2 Interactions within the community of practice

The most direct source of tutors' beliefs about and practices of assessment was the distributed cognition within the community of practice. All tutors believed they were influenced by the lecturer or senior tutors. For example, Anna mentioned the influence of lecturers on her practice of giving positive feedback:

> I think that it is, I think, I mean I haven't done a lot of educational theories, but I think it has probably changed over time. Whereas once we would have just got a mark or something you know A, B, C, D. I think that depending on who the lecturer is. I mean because I don't know there are obviously different lecturers who approach, the way that they grade and mark and do things differently. The ones that I have been lucky enough to work with have been very positive about, wanting to enhance their students learning I suppose. That's their culture. So there is a type of culture. (Anna, Interview)

121

Similarly, Cecile believed she was influenced by those who supervised her work:

> …Probably like when I first become a tutor I think like people who were teaching me or the people who were overlooking me as a tutor, probably influenced me the way of giving feedback. (Cecile, Interview)

In the community of practice, lecturers and senior tutors distributed marking schedules, sample answers, and provided other oral requirements or guidelines to tutors individually or in collective meetings (c.f. Section 4.1.3.1).

As well as the senior tutors or lecturers, the tutors were also influenced by their fellow tutors. There were various approaches to ensuring consistency among tutors who worked for the same paper at the same department. For example:

> Researcher: You mean you are learning from [[senior tutor name]]. Do you learn from others as well?
> Jo: No, Just [[senior tutor name]], maybe the other tutors as well because some other tutors are more experienced. So I guess in a sense we sort of learn from them as well. (Jo, Interview)

Mia and Emma were also influenced by tutors who worked in other departments because they were both tutoring in different departments.   For example:

> Researcher: You have that kind of meeting with the tutors and lecturers?
> Mia: Yes. Err, usually the other essays, something the other tutors might marking. (Mia, Interview)

Students were another source of tutors' beliefs. Tutors had common concerns about their students' emotional response and possible reactions to the feedback which was provided. For example, Mia tended to do "easy marking" because her "students like that" (Mia, Interview). Cecile had students come to talk with her about the feedback because "usually they think you've made a mistake or they think they should have got more mark somewhere" (Cecile, Interview). The possibility that students would be unhappy with negative feedback was the main worry for Jo who felt it was difficult to comment on the negative aspects of the assignments (Jo, Interview).

However, the distribution of information among tutors within the community of practice was uneven due to different work responsibilities tutors assumed and different interactions they had with different lecturers and students. This issue was first raised by a tutor in an open-ended question of the survey:

When completing this survey, I was thinking of a paper where the lecturer is very helpful. I have another paper where the lecturer is not helpful at all, and there is a very noticeable difference in a) my interaction with the students, b) the content of my tutorials, c) the students' ability to write good, insightful essays, and d) my ability to mark these essays. (No. 13, Open-ended comment, Survey)

The uneven distribution of information may cause contradictions within the community of practice. This point was mentioned by Mia:

…Because tutor and I are different on relevance of a section of the work. And so I had a student e-mail me, she was in my group and her friend was in the other group, she got bit flustered that they had been told one thing and we had been told another. (Mia, Interview)

The contradictions may cause some unease for tutors. Henry mentioned in the interview that once a student came and requested explanations because he had better comments on his written work from a lecturer than the comment and mark given by Henry.

…so I had a student who come up to me and he was sort of like 'Oh, I didn't quite understand why I got' – I gave him a B-, I think? And he maybe sort of thought 'I don't know why I got quite this low a mark' [mhm], um and I said um 'it's / I thought your ideas were good but you just – your writing was all, [mhm] your writing style was a bit messy, just, it was too difficult to sort of understand your ideas.' Um, he- he'd been to get a second opinion from another lecturer in the course? And then I'd marked it, and this other person has said 'Yeah, I think that seems ok'. But when I came to mark it maybe he thought // he was doing something or, you know, getting a better grade than what he actually got. (Henry, Interview)

This issue of uneven distribution of information was further raised in one of the focus group discussions:

Henry: You find a degree of anxiety is the lecturers saying something which the students, obviously take as gospel, and then, you as the tutor are expected to obviously know all that as well, but sometimes you, yeah – 'I talked to / you know, the chief lecturer and they said this um ##' // It's pretty hard to know exactly what they're on about. Or with marking, so I – and # just recently, I gave a student, maybe not that good a grade, but then he said "Oh, but I'm ## talk to [[lecturer name]] and he said [[student name]] what I was doing was fine', so like he'd already been
Mia: Like the difference between 'fine' and 'really good' and

Henry: Yeah, but then I thought 'Well' – I explained to him – I went through it and explained the marking criteria and 'This is why' in detail, but he went away okay with that ///// But just that authority thing maybe undermining your…

Frank: Or just knowing all the bits ###, I don't know, the assessment or other aspects of the course, because you'll probably get someone in the tutorial asking about@ something@.

Mia: Yeah, and if you don't know you have to go 'Well, this is what I think, but don't quote me on this' and then they'll do it EXACTLY how you think / [mm] and get it wrong.

Frank: Well, at the beginning of this paper I told my students 'I'm not some omnipotent being, I don't know everything. I'm just here to kind of, you know, give you details on the topic < . . . . > because I was usually quite good at going to the lectures' but @ past the ## haven't really, yeah, I always kind of tell them 'Look, um' – I know pretty much most of them won't, but I say 'Look, go out and do your own research, um, look up stuff on YouTube whatever, just, you know, keep yourself interested in the topic, 'cause um no matter what we tell you, there's always lots more out there and different angles on it,' so, but, yeah, a lot of them still expect you to be grand- grand [masters]

Mia: [tell them] exactly how to do to get an A. (Focus group A)

4.3.3 Reflections on personal experiences of assessment

Tutors' beliefs about assessment developed from their reflections on their own personal experiences of receiving assessment and feedback as students and providing assessment and feedback to students' written work as tutors.

4.3.3.1 Experiences of receiving assessment and feedback as students

Tutors' beliefs about feedback were also influenced by their previous learning experiences in different communities. All participants initially derived their beliefs about feedback from their own experiences of receiving feedback. For example:

Researcher: Where does this idea come from?

Emma: Partly from my own writing. Partly from / just what I've learned from classes and stuff. It is not something you are really taught at university how to write a good assignment you don't really get taught that. You just get feedback on your assignments you hand in and then you kind of develop your own writing from there. You don't really get developmental guidance as to how to write. (Emma, Interview)

Some tutors believed feedback should be positive because of their own experiences of receiving negative feedback. For example:

> One of my supervisors could be very hash, and he just wrote 'Yuk!' next to the sentences or paragraphs that he didn't like. I think it was little bit harsh so I think it probably influenced me to be more positive and try to put more positive type of feedback. I always get a lot of feedback /// I probably benefit from it. So it probably influenced me to give a lot of feedback too. So yeah I'm influenced by that. (Cecile, Interview)

Four tutors (Mia, Helen, Emma, and Eva) believed their practice was influenced by their previous academic study of English. Helen also mentioned that she had background of English study. Moreover, Helen was the only tutor who mentioned the influence of a course she took on marking and proofreading many years ago:

> Err, probably a couple of things, probably by the feedback that I have had over the years in the assignments, but also I used to work in B Block across the road /// a course on proofreading and how to mark. You know, if you were proof reading a document, what the format was to edit it, and then send it back for changes. Now I don't remember a lot of that but probably the circling and underlining is probably something that's also come from that. (Helen, Interview)

However, as mentioned by the senior tutor Jan (c.f. Section 4.3.1.3), no tutors had received teacher training. Henry mentioned that tutors could take 'training sessions' of the Teaching and Learning Development Unit in the university, but there were no systematic training for tutors on marking. Henry believed his knowledge of marking and providing feedback came from interactions with colleagues and his working experience:

> Researcher: Do you read any literature about um / giving feedback or //
> Henry: Um, no I'm not very / unless it's / I'll read all the stuff we're given [mm] / as the tutors. [yeah] and um I've been to sort of training / sessions that like by TLDU, ## [oh] um I haven't done that for maybe a little while, but every so often – particularly a lot of – if you were / relatively inexperienced or like / master's level, we'd be sent off to do like tutorial / [oh] / training sessions, just to get you know strategies for dealing with students in tutorials and that sort of thing, [oh] / um / but that was the maybe more about managing the class rather than the actual marking, [mm] I think maybe more of that comes from written stuff and / talking to other tutors and the lecturers. [Yeah] Yeah, that's where we get that kind of knowledge of giving feedback from, about actual marking / [yeah] and written comments and that sort of thing / [yeah] um, yeah, but I think the experience, yeah, so. (Henry, Interview).

Simon reflected aloud about how his second language learning experience might have influenced his actual practice of giving feedback:

> …here I'm not terribly sure – I, I might want to talk to some of the ESOL teachers or my Japanese teacher about what is a useful way of correcting people's sentences. Do you need to write the whole sentence out again, or can you just scribble on it and do people actually learn that way for second language learners, I'm not sure… (Simon, Think-aloud)

4.3.3.2 Experiences of providing assessment and feedback as tutors

Tutors' cognition of assessment and feedback developed with the increase of working experience. Four tutors (Anna, Mia, Maria and Henry) mentioned the development of their beliefs and practices over time. Mia and Maria mentioned they had become less lenient than before:

> Last year I didn't want to be too hard because it was my first time marking. But this year I have been quite clear with what they are looking for in the essay. So I kind of to be more critical. (Mia, Interview)

> I think / this semester I've become a bit more /// cynical @@@@ . (Maria, Focus group B)

By contrast, Henry believed that he had become more lenient in marking than at the beginning of his work as his contextual knowledge increased over time:

> Henry: Yeah I think I probably mark / a lot more /// um //// more leniently than maybe I used to. Maybe when I first started off I was actually quite a bit harder? Because – I don't think I was meaning to be – I wasn't trying to be mean or anything, but I think I actually // maybe I was just a bit / tougher on my marking. But now that I've done it for a long time, / Maybe I've just got more experience about how the students are? Maybe I could just see at a certain level? So if you have done a lot of first year papers you just – over several years, then you just get a feel for just the general standards for the first year students/ and then you / gradually just, you don't you know, um, / your standards / sort of balance out, in relation to what your knowledge level is, and that sort of thing, your own standards, [mm] but just what you understand of the nature of the the course and the level of the students, maybe they just came out of high school and that sort of thing, [ahh] so I think, yeah, um // a bit more um // oh easy going's not the right word, but just /// sorry, um I'm just trying to think of the right word for it.
> Researcher: More objective maybe?
> Henry: /// Yeah maybe jus- just a bit more / um / awareness of the students' sensibilities, [yeah] maybe for their age, and their

skill level. Just from experience? [yeah] yeah, so just that accumulative experience from doing things // over a long time. (Henry, Interview)

Like Henry, Anna worked out a more efficient way of assessment than before:

I mean other / years, I mean / early on I would read them all and put them into what I thought were A B C piles? But I mean it just took / so much longer, so you started to assess already just what you think and you started to er, / and then you start to put them back into different piles, and you know, so um but I @ don't do it that way anymore. (Anna, Stimulated recall)

4.3.4 Contact with members of other communities of practices

Besides the teaching and learning communities in which they were currently participating or had participated, tutors might also be influenced by their relationship with members of other communities. For example, Mia was influenced by her parents who were also working in the academic area. Mia mentioned her discussion with her friends on how to give feedback:

Researcher: How do you form this belief?
Mia: Like I said I did a lot of English classes, so I got a lot of positive feedback in English classes so I learnt how to write that way. But also my mother was a [[subject name]] lecturer. So when I first started university, she would look over my essays and tell me where to improve, so just because I was bit unsure about how to write essays at that point because my high school was bit lax, where some of the stuff they were teaching, so my mom taught me a fair bit as well. She is quite supportive in getting me to the point where I am now. And also I had a lot of friends who were doing the same kind of thing, so we would trade everything around and we would help out each other. So if we were unsure of ideas, we'd ask each other what they thought and we'd kind of get a common understanding. (Mia, Interview)

She also recalled this influence of friends and relatives after marking the assignment:

And because it is a bit hard for them to find people to read through their draft, so I just let them email to me, because I used to do the same thing when I was doing my undergraduate. I'd email my mum, and she would read my assignments for me because she was a university lecturer? (yeah) and she taught me a lot of the practice about writing university essays about not using pronouns, and um not starting with 'because' and not using rhetorical questions, not – and indenting with long quotes. She taught me a lot of those things so err, it is quite helpful to

have just someone read through it and say this doesn't make sense... (Mia, Stimulated recall)

4.3.5 Summary

This section reported the sources of the tutors' beliefs and practices. These sources included the rules and tools of assessment distributed among the tutors, the interactions between tutors, between the tutors and lecturers and students, and interactions between the tutors with people beyond the community of practice. Previous learning and working experiences also contributed to the tutors' beliefs and practices. However, no tutors had received systematic training on providing assessment feedback.

These sources of beliefs and practice are the contextual factors that influenced the tutors' beliefs and practices. These contextual factors were also sources of tutors' emotions, which is going to be reported in the following section.

4.4 Emotions and the activity of assessment

The tutors expressed their emotions in the activity of assessment and their concerns about students' emotional reactions to their assessment and feedback in both interviews and focus groups. They talked about their emotional reactions in the process of assessing students' written work during the think-aloud and stimulated recall sessions. These emotions seemed to have become integrated into their beliefs about and practices in assessment activity. These emotions will be reported in this section in the following sequence: Section 4.4.1 Tutors' emotions and beliefs; Section 4.4.2 Emotions interaction with cognition in practice.

4.4.1 Tutors' emotions and beliefs

As previously mentioned (Section 3.7.2), one of the most significant findings in this present study was the important role emotion played in the assessment activity. Tutors expressed both positive and negative emotions at work. Positive emotions included empathy, confidence, trust, enjoyment and joy; negative emotions included lack of confidence, distrust, worry or anxiety, guilt, frustration, and annoyance. These emotions derived from contextual factors that influenced their work, particularly their interactions with their colleagues and their students.

The tutors were concerned about students' emotional reactions to the outcome of assessment and they used strategies to regulate students' emotions. This section will first report the socially distributed emotions in the activity of assessment and then the emotions that interacted with cognition in practice.

4.4.1.1 Socially distributed emotion

The tutors talked about emotions that affected their work. These emotions mainly included empathy, confidence and lack of confidence, trust and distrust, pleasure, worry, frustration. These emotions had social origins in that they were derived from their interactions with their colleagues and students, or the context of their work. These emotions were also distributed among the tutors, their colleagues, and students via interaction and feedback.

Empathy

Resulting from their own experiences of learning all tutors demonstrated their understanding of students' situation:

> Because I remember when I was a student, you know in the first year, being my tutor, she didn't quite write down the whole answers and I kind of struggled a little bit, which I think if I understood those answers clearer I would have probably done better in my essay…(Mark, Interview)

> Usually I tend not to worry too much about grammar and kind of things like that. I leave that more up for them to do their essays. Some people handwrite their labs and things like that, others do it on computers. The ones who do it on computers, I tend to expect them to have better grammar. Because I know that if you make a mistake, this is what I think anyway, if you make a mistake in your handwriting and you cross it out, you like 'Oh my gosh that looks messy, I have to throw that piece of paper out and start again.' That's what I do. Er whereas if you're in a word document, you have to respect that. That's what you've got to do. Now if somebody makes a mistake, I think myself, and you know I think well perhaps they didn't have time to rewrite everything out you know and there was only one mistake, so we could probably, we won't bring it to their attention or anything like that. We would in an essay. But again this was not an essay type thing we just wanted to know what they thought. (George, Interview)

Participants all had memories of their emotional experience of receiving feedback (c.f. Section 4.1.2.3) so that they were aware of students 'emotional reactions to

positive and negative feedback.' This point will be further discussed in Section 4.4.1.2 and Section 4.4.2.

Because of their feelings of empathy, tutors tended to grade up, which seemed a strategy they used to trigger positive emotional response from students and maintain their motivation for learning.

> I tend to mark a bit higher because usually if it is a first essay, it kind of shatters confidence a little bit if they don't do really well. My students when they got their essays back they thought I'd marked them quite easily as well. They were very happy with what they got. (Mia, Interview)

This strategy was again referred to by Mia in her focus group discussion:

> The class I tutored last semester, no one knew how to write the assignment even the tutors weren't sure what to tell the students. So everyone was so like really disappointed when they got their first assignment back, no one got like an A. when I was marking because I marked a lot higher than other tutors as well. (Mia, Focus group A)

Tutors hoped their students could be encouraged to do better in the next assignment if they got a pass rather than a fail grade. However, tutors were not sure whether this was a good professional practice or not. A chain of interaction between their beliefs, emotions, and actions occurred due to the contradictory beliefs of maintaining professional standards and enhancing students' motivation:

> Henry: I'll start from the marking one that I found myself maybe more recently giving higher grades out, but, yeah I still tend to avoid trying to give out a low, like a fail, or anything like that.
> Mia. Yeah, I always get someone to read it before I want to fail it.
> Henry: I mean if it's just so bad, which sometimes they are, that you just can't help it.
> Mia: Completely awful.
> Henry: Yeah.
> Simon: Yeah, I found when I was going to fail one I'd get the lecturer to look at it but if it was really obvious, but I got really annoyed with the mark in one of my classes last year. (Focus group A)

> Frank: I think I am quite worried I was a bit too lenient while I was marking because actually I haven't failed anyone yet. But maybe I should always tell myself like I don't know like maybe if I try and get them to realize the kind of strength they can kind of work from that like give people like really low grade but never ///. (Frank, Focus group A)

130

Empathy was the only emotion which was listed among the ten topics for focus group discussion. Participants in both focus groups agreed that they felt empathetic to students. However, it also seemed to tutors that empathy was contradictory to a professional attitude towards of assessment. George even felt guilty because he graded up:

> I felt guilty about giving them a C, so I thought I'd give them a C+. (George, Interview)

In contrast to George, Henry felt guilty because he did not provide a high grade:

> Yeah and then I felt a bit guilty like I actually marked him too hard, afterwards / but // you know, that might just be 'cause you have the person there. I wasn't trying to be, you know, overtly hard or mean, anything like that at all. So it as just, and again, you get a feel for doing the whole group of assessment. (Henry, Interview)

In addition, two tutors (Maria and Helen) in focus group B mentioned the degree of empathy changed as time went by. For example:

> Maria: I think I was more – well I don't think I'd be as empathetic if I'd marked this semester as I was in the first semester. The first semester I was really, I dunno, I was really ## about what is going on, how to do it, and whether I was actually / able to say things about / how their work / so I, yeah, so if / someone had obviously TRIED, like I remember one particular one and I knew who this student was as well, which didn't help@ (@), but I knew that she tried, and that she made an effort, and she came to tutorials, and she made an effort in the tutorials, and she did the readings, and she'd written something that was actually quite well written, but she HAD NOT answered the question? [mhm] And it was AGonizing because on the one hand I was like, well she's obviously put some effort in here but she / just / – um I had to give her a lower mark, just because she
> Helen: She didn't get it.
> Maria: She didn't answer the question. She went off on her own little tangent. (Focus group B)

Confidence and lack of confidence

Confidence was specifically mentioned by six tutors. The sources of confidence included students' positive feedback and tutors' own experiences of teaching and learning. Frank gained confidence at work from students' feedback:

> …it is quite rewarding, I mean, you know, when a student kind of comes up to you and says that they really enjoyed a class you taught or whatever, you know that kind of boosts your confidence. (Frank, Interview)

May's confidence increased together with her working experience:

> Mary: I think I'll be more authoritative as well just in my own confidence like I'm much more confident about what I'm saying now. What they are saying may be not quite right /// you know, I don't want to be too sure of myself because what if they come back and they're like no [[tutor name]] I think I could have a full stop there at APA reference or something like that.
> Researcher: They did come back.
> Mary: No, no, they didn't. I was worried. So I think I would probably be more, maybe even more picky but I would just pick up more thing //// being more helpful I guess. But it is based on my own confidence as well////awareness of what I should and shouldn't do, and how to make a tutorial work, you know you've got sort of fifteen people staring at you ///
> Researcher: So you are more confident than the first semester.
> Mary: Absolutely. (Mary, Interview)

Emma had confidence at work because of her learning experience:

> This is my first semester tutoring but it is my eighth year of being a student. I have been a student a lot. (Emma, Interview)

However, new tutors seemed not confident enough at work:

> Jo: I know that some people would use like red pens which can be really very effective because it points out where your marking is. Pencil is different because we are just learning to mark, so we don't want to make some big mistake.
> Researcher: So you mean if it were not the first year for you to mark, you may use red pen.
> Jo: Yeah, maybe, if you are more confident. (Jo, Interview)

Mark, as a senior undergraduate who worked as a tutor, talked about the need for him to have confidence:

> I started getting worried and everything. But really I think, the main thing is, is just have confidence with yourself. (Mark, Interview)

Worry

It seemed that the lack of confidence resulted in another emotion, worry. This was pointed out by Mark and was also mentioned by Frank, a relatively more experienced tutor:

> Just because while it is kind of hard for someone like me because I tend to worry about things a lot like you know, 'Am I telling my students the right things,' am I err, I don't know, err

just it's kind of lot of work to make sure that you've got everything covered. (Frank, Interview)

In focus group discussion, Frank expressed his worry again:

I tend to get quite worried that I'm being a bit too lenient on people when I'm marking 'cause um I actually haven't failed anyone yet. Um, but um maybe I should have but I always kind of think to myself like you know, I don't know like maybe if I try and get them to realize their / their kind of strength, and they can kind of work from that, like I've given people like really low grades but I've never ///. (Frank, Focus group A)

Another experienced tutor, Henry, also worried about grading:

um, //// yeah I often, /// I often / worry about if / I am er / probably being a bit too harsh, to be honest, [mm] I sort of think, and am I marking this person a bit low, should I give them – you know, mark them a bit higher, or yeah, that type of thing about whether – I often agonize about the exact grade? [mm] / um, / so like this last piece of assessment, we were given a / a grading sheet, so it had the equivalent – it was out of twenty, um and something like eleven out of twenty was a C+, twelve of the twenty was B-, [mm] / um thirteen was a / B, I think, thirteen fourteen, and the A grades – or A+ was sort of seventeen onwards? [mm] but a lot of the assessments I thought 'B-', I thought quite a few were B minuses, but then when I, I actually wrote the number in, like twelve out of twenty I thought seems a little bit low, [mm] but / that was the – yeah, that was the sort of grading criteria we were given. [uh] But I did mention that to the lecturer who was the course supervisor. I said um 'Yeah, I'm just a little bit worried – I don't know whether this is a little bit low', and, he might've been 'Are you going to moderate them?' because they obviously look at the – all of the students' marks for the whole course, maybe they sort of / adjust them. [uh] Yeah. (Henry, Interview)

In addition, Henry mentioned that he worried less as he had gained more experience:

But I don't feel too worried about it – like I used to. (Henry, Interview)

Trust and distrust

Two tutors also mentioned trust between tutors and their supervisors, which may be another source of confidence:

Researcher: Will you talk about the feedback with the lecturers, or will the lecturers want to check your feedback?

Mia: They would have only wanted to check it if it was not on the same level as the other markers. That's why we checked ours first. But they were quite happy with what I was doing. They trusted me. (Mia, Interview)

Helen: I need to make sure I get the particular lecturer that I am doing what he wants me to do. His is the boss. You know, so although he's very happy for me to, I guess he seems to trust what I think and what I know. [[name]] was my lecturer a couple of years ago in a graduate paper. So obviously, he trusts where I'm coming from. (Helen, Interview)

On the other hand, the tutors mentioned that students might seek an opinion from lecturers or other tutors on their written work that was assessed by the tutors, which seems evidence of distrust. (c.f. Section 4.3.2 and Section 4.4.1.2).

Other emotions

Some other positive emotions, which could be categorized as pleasure, were also mentioned by one or two tutors:

Some of them were very, very well written. I was amazed that these people who had just came out of the high school could write so elegantly, they were really well written. (Helen, Interview)

I do really enjoy it but sometimes I'm just kind of worried that I'm not doing as good a job as I would like to. (Frank, Interview)

Frank also mentioned pleasure in his focus group discussion:

I like teaching and I like this topic pretty much. (Frank, Focus group A)

Tutors also had negative emotions, such as frustration, because they had to mark down (c.f. Section 4.1.3.2 and Section 4.2.3.2). For example, Henry felt "A little bit of frustration also a sense of sympathy for them" when he could not grade high. He even "felt guilty for marking them too hard" (Henry, Interview). Helen felt frustrated when she had to point out negative aspects of the written work:

The difference between a B+ and an A was the fact that their referencing was hopeless. It is a little bit frustrating because we had actually talked about it. (Helen, Interview)

Tutors also mentioned that other negative emotions, such as anxiety, resulted from lack of information (c.f. Section 4.3.2), agony and annoyance due to giving low

marks (c.f. empathy in this section); or getting annoyed because of the contradiction between workload and payment (c.f. Section 4.3.1.1).

In sum, tutors had both positive and negative emotions at work. These emotions were based on their beliefs and resulted from contextual factors like their interactions with other tutors, lecturers, and students at work. These emotions influenced tutors' practices.

4.4.1.2 Tutors' concerns about students' emotional reactions to feedback

Participants had common concerns about students' emotional reactions to feedback. These emotions were related to the positive and negative comments (c.f. Section 4.1.2.3). Positive feedback was believed to be able to cause positive feelings among students such as joy, confidence and willingness to study and improve. Negative feedback, whether it meant negative expressions or comments on negative aspects, was commonly believed to lead to negative feelings in students, such as annoyance, frustration, upset, embarrassment and even anger. For example, again drawing on her experience as a student, Jo believed feedback on both positive and negative aspects was effective for her but comments on negative aspects might cause negative reactions:

> Researcher: So which kind of feedback is more effective, positive or negative?
> Jo: I think they are both actually, quite good to receive, but definitely when you receive negative feedback depending on what they are talking about. It definitively makes you want to go back and think why have they said that and what have done wrong, and how can I improve on that such thing.
> Researcher: Negative feedback may push you forward?
> Jo: Yeah, But then sometimes for some people I know that it could do the opposite. I don't know at this level but I know that if you receive negative feedback and you are told you are wrong or told you don't understand, then you think 'Okay, I don't understand, and I'm going to give up.' (Jo, Interview)

Mia also believed negative feedback had negative effects:

> Researcher: Do you think you benefit more from positive feedback than negative one?
> Mia: Yes. Because if you get something positive, you think I've got something to work towards whereas if you get negative it kind of puts you off wanting to do it. (Mia, Interview)

The consideration of students' emotional response influenced how and what feedback was given. For example, both Helen and George recalled they gave feedback on the assignments written by students who had difficulty in writing. However, Helen did not give a student negative feedback on everything because she said it would be "demoralizing for him". She believed, paradoxically, that "If I mentioned everything was wrong, it would be too negative for him…But unfortunately if you did not mention it, they'll think it was OK" (Helen, Interview). However, in a similar case, George gave a lengthy feedback on everything for the student to improve even though he thought it might be annoying. "I probably wrote as much as he hated" (George, Interview). In other words, he believed it would be helpful for the student to make improvements by providing feedback on most of the negative aspects, even though the student may be unhappy to receive it. However, he tried to reduce the possible negative reaction by giving a higher mark. "I felt guilty without giving them a C; I gave them a C+. A little bit positive and I probably shouldn't" (George, Interview).

While marking, all except one participant used pencil or blue pen which was believed to be able to reduce the possible negative feelings feedback may cause. One reason of this practice was explained by George:

> But we tend to write in pencil because it's less harsh. And I think this is important. Again because you don't really want to upset people. You want them to enjoy the subject. You want them to get the subject. So you want to reward them when they get those things. But when they don't get it you don't want to put off the subject completely. If we did do that we'd probably only end up with a few people in our classes and we don't want that. We want everybody to go away thinking 'Ah! [[subject name]], I really like those papers. They were really good. I would do those papers again.' That's what we want but at the same time we do want them to learn things. (George, Interview)

Other strategies to generate positive emotional reactions including the use of ticks, and smiley faces. (c.f. 4.1.2.3).

Tutors were also aware students also had emotional reactions to grading, which may be a reason that tutors' tended to be lenient in marking:

> Researcher: Do your students come back and talk with you about the feedback?

136

Frank: Err, not often. I've had one or two that have kind of said, err, you know, 'Hey why did I get, you know, this mark?' kind of thing and usually they've been fine about it if I explain to them, err, my reasonings for, you know, giving them the feedback that I have they will be like 'Ah yeah, nah that's fair enough' which I'm probably really lucky for you know. I've known tutors who they've got students who come back to them really angry and told them, you know like, 'Why did you give me this mark? You know, I deserve much higher than that,' and they have gone into a big argument with them. (Frank, Interview)

It seems that students' emotions affected the tutors' emotions in that the tutors had empathy towards the students:

The class I was teaching last semester no one knew what – how to do the assignment, even the tutors weren't really sure what to tell the students. So everyone was so like really disappointed when they got their first assignment back and no one got like an A. I felt really bad when I was marking them because I marked a lot higher than the other tutor as well. (Mia, Focus group A)

Another reason for leniency was provided by Anna, who mentioned that students would fill in a university-mandated course appraisal form on tutors' work performance at the end each semester.

Researcher: Do your students give you feedback on your feedback?
Anna: The university, at the end, the students actually do fill out a form. Not marking. They don't go to ask you the feedback. They can if they want to. (Anna, Interview)

Therefore, students' emotional responses and reactions to feedback and assessment were a common concern among the tutors not only because of the effect that it would have on the students' motivation and improvement but may also because of its further consequences on the relationship between students and tutors, and students' appraisal of tutors' work.

4.4.2 Interaction of emotions and cognition in practice
Tutors' emotions and their concerns about their students' emotional reactions to assessment and feedback were demonstrated in their practice. Tutors talked about their emotional reactions to students' written work and marking during the think-aloud and simulated recall sessions.

The convergences between beliefs and practices regarding tutors' concerns about students' emotional reactions to assessment and feedback were found in several aspects. In think-aloud sessions, all tutors demonstrated their awareness of affective factors and tried to regulate students' emotions by a number of methods: no harsh words were used in the feedback; all except Cecile used pencil which they believed could make their feedback less harsh; ticks and encouraging words such as "Good" were used by all participants; and Mark and Cecile also addressed students by name and drew smiley faces in their feedback (c.f. Section 4.2). In stimulated recall sessions, tutors also addressed emotional factors. For example, Jo said she did not use crosses for wrong answers and tended to "mark positively" to avoid possible negative reactions.

On the other hand, it was found that tutors had emotional reactions in the process of reading, grading and providing overall comments. Three of the positive emotions- confidence, pleasure, and empathy-mentioned in the interviews were also found in the process of assessment. All the tutors also demonstrated their empathy in the process of assessment. For example:

> This is the author. They have forgot her in the reference as well, but I think it is just an oversight // because, err, / I don't know, / they might have been a little err, busy or something like that. (Mia, Think-aloud)

> Oh it is disappointing – some of them, and this particular one too actually, you know they're very clever students and they could do better work, but / to be fair, and it won't affect how I mark the assignments, they had, a lot of them had two or three assignments due on the same day / and you can see that in a couple of them who have done very well in previous assignments have obviously rushed through and thought 'I'll get something in,' you know, and um, so they're not at all up to the same standard that those particular students have produced before. I think it's just that pressure of work has built up, you know, for them, yeah, because they're mostly first year students so that you are not used to the whole – having a whole lot of things due at the same time. (Helen, Stimulated recall)

> OK. This student has gone overboard, and they have written answers here where they didn't need to, but I know who this student is and they're a very good student. So that's understandable. (Jo, Think-aloud)

Pleasure was mentioned by four participants (Emma, George, Helen, Simon). Pleasure may be because of the funny expressions in students' written work. For example, Simon laughed during the think-aloud session and then in the stimulated recall session, he explained that he laughed because it was "such a wrong word that I had laughed." Pleasure was also triggered by well written assignments. This was the case of the other three tutors who had the same emotion. For example:

> I am pleased that this one is good. (Helen, Think-aloud)
> I like the opening sentence. (Emma, Think-aloud)
>
> I love seeing work like that it just makes you feel SO happy. Then I don't have to mark them down. (George, Think-aloud)

Compared with other positive emotional reactions, confidence was not particularly evident in think-aloud sessions. In the process of reading and providing in-text feedback, Mia demonstrated her confidence in error correction because of her background of English study: she commented on an error in the assignment she marked according to her beliefs, even though it was in contradiction to the beliefs of the lecturers. However, this contradiction annoyed her:

> It's a pet peeve of mine, but apparently the school I am teaching for, the moment they're actually taught to write this way, which I've always been taught is err, / bad practice especially using err, 'I' in an essay, I've always been taught you don't – you always write about things in a detached way, rather than putting yourself into the essay and saying that 'I think that' or 'I believe that', and 'I will do this', err, I was, feel, I was always taught that that was bad practice so I tend to discourage people from doing it, err, although I haven't been doing it in this one so much because the lecturers have said that they don't mind, so these people do it so I cannot really mark them down. So doing those stuff that kind of bugs me but they know is okay. (Mia, Think-aloud)

Annoyance also emerged in the reading process:

> I find it hard to read their writing. It's too tidy and flicky which is annoying for me the marker, the curly bits on their letters. (Jo, Think-aloud)

Unlike Mia, three tutors demonstrated their lack of confidence in their assessments. For example, Helen did not provide a fail grade on an assignment, rather she decided to discuss it with the lecturer:

Yeah, it is hard. For someone like me who has only started to do this this year, yeah, I would, I'd want to discuss it with someone else. (Helen, Stimulated recall)

Here I'm not terribly sure – I I might want to talk to some of the ESOL teachers or my Japanese teacher about what is a useful way of correcting people's sentences. (Simon, Think-aloud)

Tutors felt frustrated because of the poor quality of expression of written work or because of the possibility to provide a fail grade:

…and that's the other very frustrating thing about this course // and maybe I'll call them up on it, in the final notes I do here, that a lot of Chinese students don't really go outside of their cultural boundaries. (Simon, Think-aloud)

Okay, so, not a badly argued essay, it's just / the bibliography is good, it's just hasn't referenced it properly. He's obviously used the information that he's got out of each book. And he's listed them, in his bibliography, but he just hasn't referenced them. Very frustrating. (Helen, Think-aloud)

It seems that tutors tended to express frustration when assessing students' written work. This was mentioned by Mia at the beginning of the think-aloud session before she started to mark the written assignment:

Researcher: If I wasn't in the room and you mark the assignment; do you think you would talk to that assignment?
Mia: No. Usually if I'm – if no one else is in the room I'll just like be frustrated like make frustrated noises, and then just put it down for a minute and then go back to it.
Researcher: So if I just hide behind you?
Mia: I'd still be aware of you. But no, sometimes I swear at my assignments @, but usually I'll just get frustrated and I'll put it down on a table and I'll walk away for a few minutes. (Mia, Think-aloud)

These three negative emotions, annoyance, lack of confidence, and frustration, were also found in interview data (c.f. 4.4.1.1). Other negative emotions were mentioned in the think-aloud and stimulated recall sessions. These negative emotions included disappointment, boredom, and negative reactions, all of which were because of poor quality of writing. For example:

I was thinking, that – I said it was a good opening because from the first sentence you're kind of going 'Okay, I'm interested to hear what they have to say next.' Sometimes you read an assignment and then the next – like with the next sentences I'll be going 'Okay, I'm bored already.' (Emma, Stimulated recall)

140

Phew! What! Entirely the wrong word. (Simon, Think-aloud)

Mia also mentioned that emotions changed with the increase of experience:

> So it was a little disturbing to start with and I wasn't enjoying it, but it's got better now, and I've had quite a good group of students. Um who've given me some pretty positive feedback as well. (Mia, Stimulated recall)

In sum, tutors had both negative and positive emotional responses in the process of assessment. Emotion interacted with tutors' beliefs and practice in the context of their assessment.

4.4.3 Summary

This section reported tutors' emotions during the activity of assessment. The most common emotion among the tutors was empathy, which was found in all sources of audio-recorded data. This emotion resulted from the tutors' previous experiences and affected tutors' practices of grading and providing feedback. Tutors' empathy was reflected in their concerns for students' emotional reactions and their strategies to be positive, such as avoiding a lower or fail grade, using ticks, smiley faces, and encouraging words in feedback. Other emotions such as pleasure, confidence and lack of confidence, frustration, and worry were also common emotions among tutors that interacted with both their beliefs and practices. Trust or distrust were emotions only mentioned by the tutors in interviews and focus groups, whereas some emotions, such as disappointment emerged in think-aloud and stimulated recall sessions, indicating that the conditions of students' written work caused emotional reactions. However, tutors tried to avoid passing on negative emotions but distributed positive emotions with their feedback and assessment to regulate students' emotional reactions

4.5 Summary of the findings

This chapter reported major findings of data collected by open-ended survey, interview, think-aloud, stimulated recall and focus group discussion. Four key categories were reported: beliefs, practices, sources of beliefs and practices, and emotions.

All the five sources of data demonstrated that context influenced tutors' beliefs about and practice of assessment and feedback. The contextual factors included tutors' background of learning and teaching, their interactions with other people in the community of practice of assessment, working hours, payment, and responsibility allocated to them, criteria, students' written work, and rules of assessment and giving feedback. These contextual factors interacted with tutors' beliefs and practices and caused emotional reactions. Interestingly, no tutor mentioned any influence of research literature or professional development programmes on their beliefs and practices.

Tutors' beliefs were explored by interview and focus groups. It was found that tutors had common beliefs about giving written feedback on students' assignments, such as helping students improve by feedback, focusing on content, being positive, using criteria, keeping consistency of assessment, and attending to students' affective reactions. All these factors were related to the key ingredient of the feedback, giving a grade. Whether the assessment was believed fair or not was the main concern of all those involved in writing assessment: the student, the lecturer, and the tutors. Assessment and feedback consequently influenced the development of all people involved in this practice, including the department. Therefore, good practice in giving feedback was believed to be that which could cause positive reactions of its target audience by a fair reflection of the positive and negative aspects of an assignment and a convincing justification of the assessment with professional but understandable language. However, each participant also had individual interpretations of these factors, based on their personal learning and tutoring experiences, and their knowledge of writing, assessment, students, pedagogy, and local context. This individuality of belief and working context resulted in some inconsistency and also some contradiction with their actual practices.

Tutors' practices were explored by think-aloud and stimulated recall sessions. More convergences than divergences were found between the stated beliefs and actual practices of the nine participants. Firstly, the focus of feedback was based mainly on the requirement of the written work listed in marking schedules or criteria. The amount, type, and content of feedback were limited by the contextual factors such as the quality of expression and type of written work and criteria.

Most tutors corrected grammatical errors at lexical and syntax level, partly to facilitate their understanding of the content for the purpose of marking, partly for the purpose of helping students make improvements. Secondly, tutors believed in, and used, strategies to provide positive feedback. Thirdly, tutors tried to maintain consistency of marking by various means. However, inconsistency was unavoidable because tutors had also to use their own knowledge and values while making decisions about whether to mark down the unclear expressions, whether or not to fail an assignment, and what value should be given to different aspects of assignments - especially between content, expression, and/or referencing. Tutors avoided giving fail grades because this may harm students' motivation and consequently affect their evaluation of tutors and the departments.

Divergences between beliefs and practices were mainly found in the purpose of giving feedback. Tutors believed the purpose of feedback was to help students to improve, but they mainly used feedback to justify the grades that were awarded. In addition, some participants believed they had, as undergraduates, benefited from lengthy or detailed feedback on both good and bad aspects of their assignments; however, they did not give a great deal of feedback on their own students' assignments.

The convergences and divergences between beliefs and practices indicated that tutors of the same community of practice shared common beliefs, but they adapted the beliefs to their specific context of practice.

Emotion was an important factor in their activity of assessment, which interacted with tutors' beliefs and practices. Tutors had both positive and negative emotions at work. Some of the negative and positive emotions mentioned in the interview were also evident in the think-aloud and stimulated recall sessions. Some other negative emotional reactions were found in the process of think-aloud. Tutors had negative emotional responses to grading when they had to choose between maintaining professional standards and being lenient in marking in order to maintain students' motivation. Tutors were also aware of students' emotional reaction and response to assessment and feedback, therefore they tried to regulate students' possible negative emotions by providing positive feedback or grading

143

up. These findings provided insight to the research questions, which will be discussed in the following chapter.

CHAPTER FIVE: DISCUSSION

5.0 Introduction

This research project is expected to contribute to the theoretical framework of teacher cognition by exploring the convergences and divergences of the beliefs about and practices of subject tutors in the activity of assessing writing and giving feedback, both as individuals and as members of specific communities of practice in a New Zealand university.

As described in Chapter Three, data were collected by multiple methods and then analyzed first by NVivo8 and then manually. In Chapter Four, the following categories were discussed regarding these tutors' beliefs about and practices of assessing academic writing from various disciplines: good writing, feedback, grading, criteria and, crucially, the importance of emotional and contextual factors in the activity of assessment. The main divergences between tutors' beliefs and practices were their stated purpose of feedback (improvement of learning) and the actual purpose served in practice (justification of grade), and their beliefs and practices in regard to error correction.

The findings (summarized in Section 5.1) will be discussed in this chapter, to address the primary interest of the research. This discussion will relate the findings of the study to the literature covered in Chapter Two, with particular reference to the issues which emerged in the findings that had not been fully explored in previous studies. These issues are: the significant roles that emotion and context play in the tutors' cognition; and practices in the activity of assessment. The emergence of the two issues has been facilitated by the multiple method approach, especially the use of think aloud and stimulated recall, the implications of which will be discussed later in the chapter.

The first part of the discussion (Section 5.2) is the theoretical framework that overarches the discussion. The starting point is the role that context plays in individuals' cognitive development and the language mediated internalization and externalization of individual cognition (Vygotsky, 1978, 1981). The context of the tutors' beliefs and practices will be analysed first, with the help of

Engeström's (1987) expanded model of activity theory. The interactions and contradictions between individual cognition and contextual factors which can cause both emotional and cognitive reactions will be analysed. The causal relationship between emotions and cognitions will be analysed according to the concepts of cognitive dissonance (Festinger, 1957) and cognitive reappraisal (Gross, 2002, 2008). The second part of discussion (Section 5.3) is the analysis of the internalization and externalization processes, which include both cognitive and emotional experiences of individual tutors. This will analyse how emotion, cognition, and action (ECA) interact with each other in the process of assessing writing by individual tutors. The third part of the discussion (Section 5.4), will be an analysis of the distributive nature of cognition and emotion and how this affects the practices of individual tutors. The fourth part of discussion explains how the individual tutors' ECA is regulated in the context of the activity of assessment, the contradictions that need to be identified, and support that is needed. It will be argued that contradictions cause cognitive dissonance and emotional reactions which strongly influence cognitive development and actions within activities such as providing feedback on student's written assignments. This will explain the causal relationship between contradictions of the activity system, individual and distributed cognition, emotion, and the observable actions, thus to provide an understanding of the factors that influence tutors' cognitive development. The conclusion (Section 5.6) will argue that the tutors' personal emotional experience is part of their cognitive condition. The tutors' emotional reactions are triggered by cognitive dissonance and consonance between individually and socially distributed cognition. The contradictions within the current activity system result in negative emotions, which affect the outcome of the activity. A new model of cognitive distribution that is supported by the activity system is needed to enhance information flow and cognitive consonance within the activity system so as to achieve the goal of collective cognitive transformation by expansive learning.

The structure of this chapter is: 5.1 Summary of findings; 5.2 Context of the individual and collective cognition and emotion of the activity of assessing disciplinary writing; 5.3 Language mediated interactions between cognition and emotion in the process of individual tutors providing assessment feedback; 5.4 ECA interactions at the collective level of the activity of providing assessment

feedback; 5.5 Regulation of ECA and contradictions in the activity system; 5.6 Conclusion.

5.1 Summary of findings

The study found that tutors had convergent and divergent beliefs about and practices of assessment which were related to the sources of their beliefs. Emotion was found to influence tutors' beliefs and practices in regard to assessment.

There were more convergences than divergences among tutors in regard to their beliefs. Divergent beliefs among tutors were mainly found regarding how to use the criteria; how many errors should be commented on and whether errors should be marked down; and how to be positive in assessment feedback. Convergently, tutors believed that they assessed written work according to criteria or marking schedules provided by those who supervised their work but they also used their personal knowledge of writing and assessment. Another convergence was that tutors believed the purpose of giving assessment feedback was to help students make improvements. Tutors also believed that good written work should be clear. They said that they focused on content when grading but also provided feedback on grammatical errors. Their major concerns about assessment were being consistent in their marking and being fair to all students. More interestingly, tutors believed that students might have emotional reactions to grades and feedback; therefore, tutors tried to be positive when assessing students' written work and providing feedback. Finally, tutors were emotionally engaged with the activity of assessment – a major finding which will be discussed in detail later in this chapter.

Convergent practices were found in four main aspects. Firstly, tutors had common routine processes of marking: reading the written work and providing in-text feedback, weighing different aspects of the written work, composing overall comments, and calculating grades. Secondly, the most common strategy tutors used in assessment was making comparisons between marked assignments to maintain consistency. Thirdly, tutors had a common tendency to mark higher when marks fell between two scales. Frequently, the process of marking triggered tutors' emotional reactions. Compared with the common routines, more divergences were found in the detailed process of assessment among tutors.

Tutors used criteria or marking schedules differently in the process of marking; they focused on different aspects of the written work being assessed; they also demonstrated different ways of being positive; in addition, the average time tutors spent on assessment and the length of overall comment provided by different tutors varied.

A comparison between tutors' beliefs and practice demonstrated three major contradictions. Firstly, all tutors believed that content was the focus of their feedback, but the in-text feedback for those who marked essays or reviews was mainly on grammatical issues. Secondly, some tutors believed they gave plenty of feedback but this was not observed in their practice. Thirdly, all tutors believed that the purpose of giving feedback was to help students to improve, but in practice, they rarely provided suggestions on how to improve. The study found that the convergences and divergences among tutors' beliefs and practices related to the sources of tutors' beliefs. The most common source of tutors' beliefs was the lecturers and/or senior tutors supervising their work who provided oral instructions, marking schedules, guidelines, and sample answers. The participants in this present study also exchanged opinions with some other tutors in the same faculty. Other sources included the feedback and grade tutors had received on their own assignments as students, and their own experiences of assessment and giving feedback. In addition, relevant knowledge from previous learning experiences in general, and friends or family members, could also influence tutors' beliefs. However, no tutor received formal training on academic writing, teaching, or assessment; nor was there evidence that tutors were influenced by the literature of writing assessment or feedback research. Tutors' practices were influenced not only by their beliefs but also by institutional regulations as well as dynamic contextual factors such as the quality of expression of assignment, the usability of criteria, time limit, and the physical conditions of tutors such as tiredness.

Most importantly, emotion was found to be an important ingredient in all sources of recorded data. Tutors not only expressed their concerns about students' emotional aspects of assessment but also expressed and demonstrated that emotions such as confidence, empathy, and trust influenced their practices. Three major roles of emotion were found in the present study. Firstly, the tutors had memories of their emotional reactions to assessment feedback on their own

written work. Secondly, the tutors had a wide range of emotions from interactions with their colleagues and students in the assessment activity. Thirdly, tutors had emotional reactions in the process of providing assessment feedback. These findings will be discussed in the following sections from the perspective of activity theory.

5.2 Context of the individual and collective cognition and emotion

A Vygotskian socio-historical perspective of cognition regards context as the key to the understanding of cognitive development. Context not only constrains but also affords cognitive development (Vygotsky, 1987). The context of tutors' cognitive development of assessment activity in this study is the activity system within the university, which can be analysed from a contextualization of Engeström's (1987) expanded model of the activity system (Figure 5.1).

Artefacts such as language, guidelines
marking schedules, and criteria and
physical tools such as pencils and pens

Intended
↗ outcome
Written work →Actualized
↕ outcome
Undergraduate

Tutors

Rules and conventions of
assessing specific
assignments, time,
payment; General policies
of assessment and grading,
the types of assignment
considered academically
appropriate, and
conventions of interactions

Division of
labour

Community composed
of professors, lecturers,
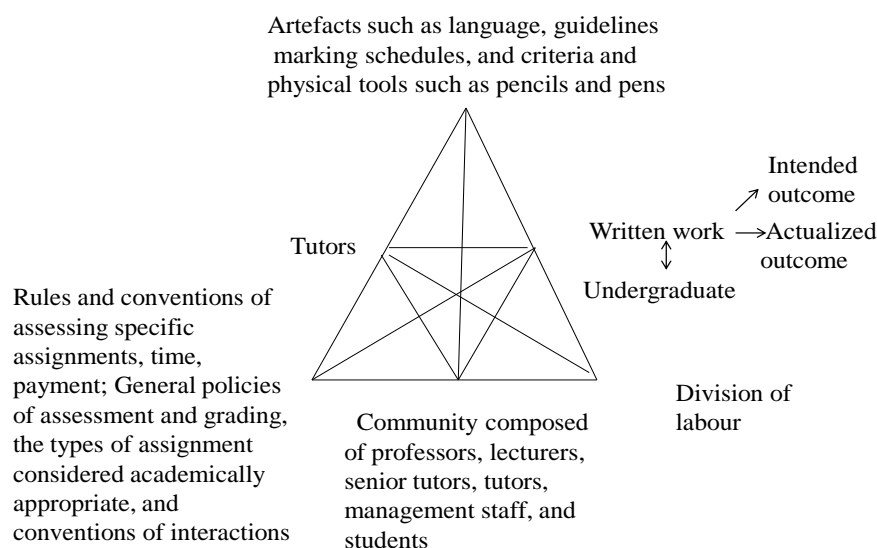senior tutors, tutors,
management staff, and
students

Figure 5.1: Expanded model of activity theory (adapted from Engeström, 1987, p. 78)

As illustrated in Figure 5.1, the contextual elements of the activity of assessing disciplinary writing in the present study included several elements. The prior condition of the activity was the motivation: to provide fair and consistent evaluation of students' learning and to inform students of their learning to enhance further achievement. The intended outcome in this case, was the goal of students' cognitive improvement; and the actualised outcome, the summative

assessment and grades. The second component was the individual tutors who assess undergraduates' written work. The third component was the written work of undergraduate students as the object of the activity. The fourth component was the tools of assessment, including both material tools, such as pencils, pens, and computers, and cultural artefacts, such as the criteria, marking schedules, guidelines according to which students' written work was assessed, and written language used to compose assessment feedback. The fifth component was the community in the activity system, which was composed of professors, lecturers and senior tutors, management staff, and students. Among the community, those who assessed the students' written work formed a hierarchical structure of community of practice with novice tutors at the peripheral edge of the community due to the division of labour. This community of practice followed the rules established by the management groups. The roles students played in the community only included two aspects: completing and submitting their assignments, and providing an appraisal form at the end of the course. The sixth component was the division of labour, including the horizontal division of tasks and the vertical division of power relationship and status among the members of the community. The last component was the rules of assessment, which included the institutional policy of assessment and grading, the explicit rules or implicit conventions of assessing assignments of specific courses, the types of assessment considered academically appropriate, the conventions of interactions within the community, and rules of payment.

These contextual elements interacted with each other and regulated individual tutors' beliefs and practices. The mechanism of regulation, according to Vygotsky (1978), included both other-regulation and self-regulation. Other-regulation in this present study involved two aspects. Firstly, the individual tutors were affected by other people in the activity system, and mediated by cultural artefacts such as written documents and oral interactions; secondly, the tutors regulated the students' cognitive development, using the assessment feedback as a mediating tool. Self-regulation also had two aspects: individuals internalized, via inner speech, understanding reached on the social plane, and then externalized their assessment decisions, using the medium of written feedback. However, it seems that Vygotsky (1978) did not explain the mechanism of the interaction between emotion, cognition, and action. The causal relationship between emotion and

cognition has been identified by Festinger (1957) using the concept of cognitive dissonance. According to Festinger and Carlsmith (1959), people suffer cognitive dissonance when they hold contradictory concepts or when they face a choice between contradictory attitudes and behaviours, which in turn triggers negative emotions. According to later developments in psychology, cognitive reappraisal of the situation can reduce emotional reaction and regulate actions in order to achieve the intended goal (Gross, 2002, 2008). The following diagram that is based on my understanding of Vygotsky's (1978) principle of internalization and externalization can interpret the process of the ECA interaction of individual tutors:
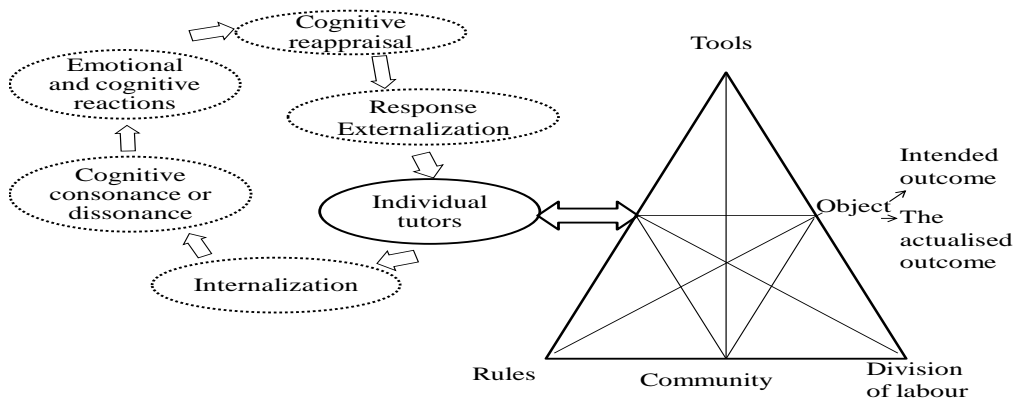

Figure 5.2: Process of ECA interaction in the activity of assessment

In the activity of assessment, the individual tutors received information about contextual factors from the activity system. The tutors internalized this contextual information and compared it with their existing cognitive framework, which resulted in cognitive consonance or dissonance. Cognitive consonance might cause positive emotional reactions such as happiness or enjoyment; cognitive dissonance might cause negative emotional reactions such as frustration or annoyance. Cognitive reappraisal of the contextual factors according to the goals could help the tutors decide on the extent of externalization of their emotions and cognition. The result of externalization was the assessment feedback on the written scripts.

In sum, the tutors were not only cognitively but also emotionally involved in the assessment activity. They tried to distribute positive emotions and avoid

distributing negative emotions. An analysis of the ECA interactions contributes to the understanding of Vygotsky's (1978) concept of internalization and externalization.

Context has long been regarded as crucial to the understanding of convergences and divergences of teachers' beliefs and practices (Basturkmen et al., 2004; Crookes & Arakaki, 1999; Lee, 2009). However, the problem of the existing constructs of context in the relevant literature is that they are either too vague or too general to be analysed in specific studies or they are just physically or geographically defined as "around and inside the classroom" (Borg, 2006, p. 283) which is too narrow a perspective to reveal the complex interactions between contextual factors. Cross (2010) applied activity theory to his analysis of context. However, he did not use think-aloud in his study and did not analyse the ongoing interactions between contextual factors and the individuals' ECA. According to Engeström (2009), there is a need to address individual and emotional issues in activity theory. The present study occupies an important research space in teacher cognition studies by applying activity theory to the analysis of context and it contributes to the development of activity theory by its analysis of the ECA interactions of individuals.

The unit of analysis in the present study is the activity of assessment that includes the contextual factors that affect the beliefs and practices of both individuals and the community of the activity to which these individuals belong. In analysing the beliefs and practices of both individual and the communities, the concept of community of practice according to situated cognition will be applied, and there will be an analysis of the relationship between tutors as a group of agents and the system that mediates their cooperating action, and which concepts were also the focus of distributed cognition (Nardi, 1996). The concepts of cognitive consonance and cognitive appraisal to the analysis of the ECA interactions will be applied.

The following sections will analyse tutor cognition from three aspects: the ECA interactions of individuals, the ECA interactions between individuals, and organizational or structural regulation of cognition of individuals.

5.3 Language-mediated interactions between cognition and emotion in the process of individual tutors providing assessment feedback

The findings of the present study revealed that language plays two complementary roles in the process of internalization and externalization of cognition and emotion. On one hand, language mediated cognition and it did so in the following ways. First, the language used by the students in their assignments mediated tutors' understanding of the written work. The tutors all believed good written work should be clear because language, as the mediator of cognition, affected their understanding of content. The tutors had to give feedback on grammatical errors to be able to subsequently focus on content. Then, language in the form of written documents such as marking schedules and oral instructions on assessment was a crucial tool for tutors to assess students' written work. Finally, as the product of the tutors' work, language mediated cognitive development of students, although the tutors were not sure of the mediating effect of their feedback in their students' cognition. On the other hand, language also mediated emotion. The present study found that the tutors had emotional reactions to students' written work. The tutors had positive emotions when they found the quality of students' work was consistent with their own expectations and/or the formal requirements of the course. By contrast, negative emotions emerged when they confronted cognitive dissonance. The tutors felt "frustrated", or "guilty", or "bad", or "awful" when they had to give a lower or a fail grade because a low or fail grade was in contrast to their beliefs about being positive. They also felt guilty or worried if they gave relatively higher marks to encourage students which they felt might jeopardize professional standards of fairness. However, some strategies tutors used might have related to their worry about students' possible negative reactions to the assessment. For example, some tutors used several rounds of marking or making comparisons between different written scripts to ensure the consistency of grading. Moreover, the tutors used feedback to regulate students' emotional reactions because they appreciated the potential impacts that emotions have in students' subsequent actions. To reduce students' negative emotion, tutors tried to distribute positive emotional signals to students by smiley face icons, ticks, encouraging words, as well as feedback on positive aspects of the assignments.

The processes of internalization and externalization and the interactions between ECA were found in the present study during think-aloud sessions and were explored in the subsequent stimulated recall sessions. The tutors internalized the contextual factors and externalized their cognition by assessment feedback which was intended to regulate students' cognitive and emotional responses. The interactions between ECA can be analysed with the help of Leont'ev's (1978) the three-level activity structure. According to Leont'ev (1978), the collective activity is composed of goal-oriented actions and condition-dependent operations. Routine operations, which are carried out automatically or subconsciously, vary according to different conditions. All operations initially come into existence as actions because they challenge the existing cognitive conditions, and thus require conscious deliberation (Section 2.4.4.1). The tutors who participated in the activity of assessment had routine operations in the process of assessment, depending on the conditions of the type (essays and lab reports) and quality (clarity and content) of the written scripts, type and content of criteria or marking schedules, and requirements. The conditions triggered different interactions between ECA of individual tutors. These low level operations turned into actions when extra cognitive efforts were required to solve problems. This section will discuss the ECA interaction of the tutors, firstly in the process of reading and providing in-text feedback, and then in the process of allocating grades.

5.3.1 ECA interaction in the process of reading and writing in-text feedback

The first part of the assessment process of the tutors was reading, and then writing in-text feedback. During this process, clarity of the written scripts played an important role as tutors had to understand the expressions before evaluating the content. They regarded this as the most important feature of good written scripts. The main threat to clarity was perceived as grammatical errors at lexical and syntax level; and it is important to note that the tutors who assessed essays felt they needed to correct the errors before they could focus on the content of the written scripts. This raises the issue of whether the tutors would be able to focus on the content without attending to grammatical errors, especially given the time they were expected to spend on assessing a piece of written work.

Essays were more complex than lab reports in expression (grammar, organization, references, syntax, and lexical choice) and the criteria for essays were usually holistic. The tutors had to interpret both the criteria and the written scripts according to their existing cognition of assessing writing which had been developed historically through their participation in relevant activities. Different sources of beliefs resulted into different operations. However, the clarity of the written scripts determined whether or not the existing cognitive conditions were challenged in the process of reading. While reading essays with few grammatical errors, little or no extra cognitive effort was needed in understanding the content. Therefore, the operations were reading while identifying and correcting errors and providing brief in-text feedback regarding organization and/or content (such as ticks for good points). However, extra cognitive effort was required to understand the written scripts with too many errors, especially errors committed by non-native English speakers. In this condition, the operations of reading rose to the level of actions. As tutors' cognitive conditions varied, so did the actions they took: some tutors who were confident in error correction usually picked out most of the errors, and corrected or gave indirect feedback on the errors. In a case where there were expressions that they could hardly understand, tutors usually highlighted the expressions by symbols (e.g. underlining) and spent a longer time rereading and sometimes rephrasing the sentences to ascertain the meanings. Tutors who had less confidence in error correction tried to work out meanings by taking longer time to reread the written scripts while perhaps correcting some of the errors.

Grammatical errors in lab reports were usually less of a challenge to tutors' existing cognitive conditions than those in essays because lab reports were usually composed of brief right or wrong answers. This meant that tutors were able to make judgments based on analytical marking schedules. In this condition, the operations of reading lab reports were more a process of looking for correct answers mediated by the analytical marking schedules, including comparing students' answers with those given in the marking schedules, ticking the correct answers, crossing the wrong answers and sometimes providing correct answers. According to Barkaoui (2007), analytical criteria generate more interaction between raters and scales, while holistic criteria generate more comparison

between written scripts. The present study seems support Barkaoui's (2007) finding.

Regardless of the type of written scripts, high quality written scripts (expression and content) triggered positive emotional reactions among the tutors. These reactions were based on cognitive appraisal of the possible outcome of assessment that a well written script meant a high grade which would satisfy the expectations of both students and the institution and, in turn, fulfil tutors' motivation to participate in the assessment activity. In addition, well written scripts saved tutors time and cognitive effort in reading and providing corrective feedback and justifying a grade. By contrast, written scripts which had more grammatical errors caused negative emotional reactions such as frustration and annoyance during the process of reading. Although sometimes tutors laughed at some strange ways of expression in the written scripts, they basically did not enjoy reading written scripts which were full of grammatical errors because it increased their cognitive load and took them longer time to read.

Tutors demonstrated certain regulatory control over their emotional reactions to the written scripts. Tutors would like to indicate their positive emotional response to the written scripts by using double ticks and praising words. However, the negative emotions which emerged in the process of reading were not found directly reflected in feedback except that more in-text comments including symbols (such as underlining, crosses, question marks, etc.) were used. Although some tutors believed that too much in-text feedback on negative aspects of writing might be "too negative" for students, they still did so because they believed it would be beneficial for students.

The detailed analysis of ECA of error correction shifting between operations and actions within the activity of assessment not only contributes to an understanding of the divergence between tutors' beliefs and practices regarding the purpose of assessment, but also contributes to the existing academic understanding of feedback. This study confirms that the tutors' actualised outcome of error correction was assessment rather than the intended outcome of learning improvement. Therefore, the extent of error correction was usually an optional choice in the activity of assessment based on tutors' willingness, competence,

156

time limits, and sometimes oral instructions from their supervisors on whether or not to correct errors. Moreover, as noted above, the correction of errors enabled the tutors to make sense of ill-written scripts, and thus could be considered a rational strategy to regulate the tutor's actions (even if of little benefit to the students). Further research is needed to explore the reasons given by markers in other contexts who do focus on lower order concerns. There was an absence of knowledge or strategies on issues of how to correct errors, how much to correct, and specific goals to be achieved for individual students by so doing.

The explanation of why errors were corrected or not in these tutors' practices of assessing disciplinary writing contributes to the discussion in applied linguistics circles regarding the effectiveness of error correction (e.g. Ferris, 1999, 2006; Truscott, 1999, 2007). Error correction has been found to be effective in some empirical studies (e.g. Bitchener, 2008; Chandler, 2003; Ferris, 1999, 2006) but this is usually based on continuous feedback on a limited number of specific aspects of grammatical errors identified in students' compositions. The effectiveness of error correction seems merely an assumption when, as in the present case, no continuity of feedback on the errors could be ensured due to the modular pattern of courses where feedback was only provided on the final version of written work, and not on drafts. Most of the tutors in the present case were not confident of their ability to correct errors effectively, an issue which had been raised by Truscott (1996, 1999) in his argument against error correction.

This study also contributes to the discussion on the assessment of disciplinary writing regarding feedback on language. In addition, it contributes to the understanding of the current practice of feedback on discipline features of writing. According to Carter (2007), there has been an increasing awareness on the specific discourse features of discipline writing among teachers. This seemed true in this study when some tutors talked about the features of writing in different disciplines, although there was almost no in-text feedback given by tutors on discourse features of academic writing.

5.3.2 ECA interaction in the process of grading and drafting overall comment

As has been mentioned above, the assessment of lab reports was usually not a challenge to tutors' existing cognition in that they could find right or wrong answers to most questions (if not all) in the analytical marking schedules. The result of assessment was based mainly on the conditions of written scripts and the marking schedules. Tutors' cognitive condition was hardly challenged as the grading process was merely the operation of calculating marks. Thus, as far as lab reports are concerned, few contradictions between ECA were observed.

By contrast, more contradictory interactions of ECA were found in the process of grading essays because the criteria did not specify right or wrong answers. Usually, the following actions were taken by the tutors when grading essays: weighing each aspect of the written scripts according to the requirements listed in the criteria; making comparisons between written scripts of the same cohort students; and awarding and perhaps adjusting the grades when necessary. The goal of taking these actions was to maintain consistency and fairness of assessment. According to studies in the assessment of second language writing, inconsistency of assessment results mainly from raters' various backgrounds (Hamp-Lyons, 1991; Sakyi, 2000; Vaughan, 1991; Weigle, 1999). However, the inconsistency of grading between tutors found in this study resulted from their different beliefs on how much emphasis to put on different aspects of writing while weighting the grades, including whether or not to mark errors down, and different degrees of empathy they had toward students. The inconsistency among the tutors' cognitive conditions could upscale the actions of grading into activities that required interaction between tutors and other members of the community of practice, especially when students received a low or fail grade from their tutor yet a better evaluation from another tutor or lecturer. Therefore, tutors usually avoided giving a fail grade and sometimes preferred to seek advice from other tutors or lecturers on whether or not to award a fail grade, in which case, the individual action went up to the level of activity. Compared with the experienced teachers who perceived themselves as judges in Jeffery and Selting's (1999) study, tutors in this study were aware of their identity as being less authoritative than lecturers in making assessment decisions. Additionally, those tutors in this study

who had less experience of assessment activity were less confident in making appropriate judgments than the relatively more experienced tutors.

With regard to emotional issues, the tutors felt frustrated when a written script did not meet the requirement of a pass grade. They were aware that inconsistency of grading and low marks might well cause negative emotional reactions, such as anger and annoyance, on the part of students. They tried to balance the evaluative function of assessment and its possible negative emotional impact on students for fear that it might have the consequence of affecting students' engagement with the assessment feedback, motivation to learn, and evaluation of the course. Therefore, their course of action tended to be lenient when marks fell between scales. However, some tutors also worried whether they were so lenient as to make their assessment unreliable. It seemed that the degree of empathy towards students and worry about work performance reduced as experience in assessment increased.

Another action taken by most of the tutors was to provide an overall comment to students at the end of the written scripts or on the feedback sheet. This action was taken either before or after the grading process, mainly for the purpose of justifying the grade to both students and those who supervised the tutors' work. This purpose could be found in the following aspects: firstly, the form of the overall comment was usually a summary of what had been done well and what had not been fulfilled according to the marking schedule or criteria. In the think-aloud sessions, the tutors were usually silent when drafting the overall comment, suggesting that complex cognitive processing was needed. Two tutors who assessed essays would provide an overall comment later after comparison with some other written scripts. Secondly, the general length of the overall comment on essays was much longer than those of lab reports. Thirdly, like the in-text feedback, no information about how to improve was provided in the overall comments except that some tutors mentioned they might suggest students go to the learning support unit in the university to improve their grammar.

It seemed that tutors also intended to use overall comments to enhance students' positive emotional reactions to assessment and/or reduce the negative ones. This purpose was found in the process of drafting overall comments: tutors who assessed well written essays tended to express their satisfaction and pleasure in the overall comment by praise, while in less well written scripts, tutors tried to be

positive by identifying positive aspects at the beginning of the overall comment, following these with the negative aspects. Some tutors selected appropriate encouraging words; some chose a conversational style, starting the overall comment by addressing the student by name and/or finishing with a smiley face icon and/or the tutor's name.

To sum up, emotional reactions emerged during the process of assessing students' writing. Positive emotions such as pleasure and enjoyment resulted from high quality written work. These positive emotional reactions were intentionally regulated: a response was cognitively formulated and then distributed to students by the written feedback. On the other hand, negative emotions such as frustration and annoyance were not reflected directly in the overall comments. Tutors tried to reduce possibly negative emotional reactions of students by encouraging words or lenient grading. Therefore, the process of assessing writing in a sense was a double process of regulation of emotion and action: firstly, the tutors internally regulated their emotional reactions, and secondly they regulated the students' emotional reactions by their feedback to students, which was a form of other-regulation. The result of regulation varied due to tutors' divergent personal backgrounds and other contextual elements.

The findings and analysis of ECA contributes to the existing literature related to the interactions between cognition and practice in the process of assessment in university context. Existing studies have explored the interactions between raters' background experience and their marking process (Vaughan, 1991) between written scripts and raters' expectations of writing in reading process (Huot, 1993), and between criteria and marking process (Barkaoui, 2007). However, these studies on rater cognition were usually carried out in controlled experimental conditions focusing on cognitive processing that resulted in assessment proficiency. Among the very limited studies in natural contexts of assessment, the one carried out by Jeffery and Selting (1999) focused on the perceived identity of assessors in the process of providing feedback. Contextual and emotional issues have not been explored in any of these studies. As far as I can identify, the study by Crisp (2008) is the only one that has started to explore the relationship between personal, emotional, and social factors and marking process. However, Crisp did not specifically categorise the emotions found in her study and did not analyse in

detail how emotion interacted with cognition. The interactive factors discussed in these studies were found to play a role in the assessment activity in this study. However, more contextual elements were found in this study in the process of assessment that interacted with tutor cognition and emotion at both individual and social levels.

5.3.3 Summary

Based on Vygotsky's (1978) concept of internalization and externalization and Leont'ev's (1978) three levels analysis of activity, this section discussed tutors' self-regulation of ECA in the process of assessing students' written work, firstly in regard to the processes of reading and providing in-text feedback, and then the process of grade-weighting and providing overall comments. It explained that the quality of expression of the written scripts triggered tutors' emotional reactions. Poor quality of expression resulted in error correction, which attracted the tutors' attention to the lower level concerns at the expense of making decisions at higher level of writing. Actions were taken in the process of weighting and deciding on a grade, especially a fail grade. It was explained that the tutors' confidence played a role in both error correction and grading. The discussion in this section also reveals the summative goal of assessment in regulating the tutors in their practice. The ECA interaction discussed in this section was embedded in the context of the activity where the tutors' cognition and emotion were historically developed, which will be the focus of discussion in the next section.

5.4 ECA interactions at the collective level of the activity of providing assessment feedback

The present study found that the tutors' cognitive experiences were accompanied with emotional experiences. The ECA interactions happened not only when tutors undertook operations or actions in the process of assessing writing but also occurred between the tutors and their colleagues and students at the level of collective activity. This finding is in alignment with previous studies (e.g. Lazarus, 1991; Roth, 2007; Sutton &Wheatley, 2003) that social activity generates socially distributed cognition and emotions. Moreover, the present study found that assessment feedback mediated emotions between the tutors and the students. Positive emotions came from the response or feedback from others that

value the effort of their work or the achievement of the goal; whereas negative emotions, such as worry and frustration, emerged when the there was no recognition of the effort of achievement of the goal. The interactions between ECA in the activity of assessment will be further analysed in the following sections in regard to three emotions: empathy, confidence, and trust.

5.4.1 Empathy and being positive in assessment

Tutors had empathy towards students. This emotion affected their decision-making in the assessment activity and, in turn, regulated students' emotional reaction to the grade and feedback. This finding is in contrast to the conclusion made by Varlander (2008) who argued that it was not easy for tutors to have empathy towards students because they were in a different situation; thus they needed to train themselves to be empathetic. The reason, perhaps, is the dual identity of tutors in this study which is different from those of experienced teachers in other studies. The main source of empathy was tutors' own memories (some very recent) of emotional reactions when receiving assessment feedback, which included pleasure on receiving high grades, and anger, annoyance, or embarrassment on receiving low grades or negative feedback. For example, one tutor in the interview said she felt it was too harsh when there were negative words such as "Yuk!" in the feedback of her own assignments. Another tutor felt offended when he read a negative comment on his assignment regarding his language ability. Tutors drew upon their emotional experience to their current work and tried to avoid using negative feedback that could cause negative emotions among students. This finding concurs with the findings of previous studies that negative feedback triggers negative emotions (Trope et al., 2001), affects their motivation (e.g. Ferris, 1995; Ivanic et al., 2000), attitudes (Storms & Sheingold, 1999), and their self-esteem (Ivanic et al., 2000) .

The tutors were aware of the positive function of empathy as a regulating tool to reduce the possible negative effects of assessment, and possibly to generate positive emotional reactions by students. All tutors except one used pencils or blue pens instead of red pens to mark the written work because they believed pencil marking seemed 'softer' than red-pen marking. Tutors were aware that the use of negative expressions or symbols such as crosses could cause negative

162

emotions. The reasons for this practice are, perhaps, that positive emotional reaction to the result of assessment, on one hand, is consistent with the fulfilment of the summative goal of assessment, and on the other hand, could motivate students to make further improvement. However, the tutors had divergent beliefs on how to be positive. This indicates that tutors had tacit knowledge from their experience but there were no explicit theories or rules to guide them on how to mediate students' emotion and cognition by assessment feedback.

The tutors tended to be lenient towards students when assessing their written work: they tended to grade up when there was possibility of providing a bare pass or fail grade, or they corrected but did not mark down errors. Tutors believed that a higher grade as an encouragement is more effective to motivate students in their study than a fail grade calculated according to the product of the written work against the criteria. It should be noted that being lenient in grading seemed also to be interpreted by tutors as one aspect of the "be positive policy" informally distributed among tutors. In fact, being positive is the main strategy shared by tutors to facilitate their commonly believed formative role of assessment feedback. This finding seems to support the finding of a previous study by Hawe (2003) that lecturers were reluctant to provide a fail grade in a New Zealand college of education. According to Hawe, grading is not only an evaluation of a piece of work, it is always personalized and related to the success or failure of a student. Moreover, there is institutional pressure against fail grades due to funding and employment concerns. Being supportive and positive in weighting grades is also one aspect of being fair to students and being a good teacher.

However, tutors had different levels of empathy according to their personal experiences. One tutor, in a focus group discussion, mentioned that she became less empathetic in the second semester than in the first semester of her work because of her increased experience. By contrast, however, another tutor mentioned he became more empathetic towards students' writing after years of experience because of his increased understanding of contextual issues such as students' ability to write. One common reason for the change of both tutors was their cognitive improvement, although such enhancement resulted in opposite changes of empathy. Although tutors were empathetic and consciously applied this empathy towards students in assessment to regulate students' emotions, they

163

were not sure whether empathy was acceptable professional practice or not. Some of them worried whether this made them too lenient to students. It seemed to the tutors that empathy might be in contradiction to professional standards of assessment.

5.4.2 Confidence in carrying out the activity of assessment

Confidence was an emotion mentioned by most of the interviewed tutors and demonstrated by all tutors in think-aloud and stimulated recall sessions. As has been discussed in Section 5.3, confidence interacted with the tutors' emotions and cognition in the process of assessing writing. Tutors who had not specialized in English had less confidence in error correction than those who had. Furthermore, tutors felt less confident and more hesitant when deciding on whether or not to award a fail grade; sometimes they wanted to seek advice from their supervisors. Confidence interacted with the tutors' existing knowledge at work, influenced the decision making process and resulted in tutors' different operations or actions. This finding concurs with Roth (2007) that emotions influence decision-making.

New tutors, due to the lack of experience, did not have enough confidence. Confidence also seemed to be related to the status of the tutors in the community of practice. For example, the tutor who was a senior undergraduate talked about his lack of confidence because the other tutors were master's students. The lack of confidence made the tutors worry about their performance. Tutors' confidence increased when they became more experienced or when they received positive responses to their work from students and those who supervised their work. This indicated that confidence not only resulted from individual experiences and qualifications but also was related to other people's evaluation of their work.

Confidence in assessing writing is an emotion that has been investigated by Goos and Hughes (2010). They found, by an on-line survey, that course coordinators were more confident in making assessment judgment but had less confidence in the institutional requirements for their assessment performance. Compared with Goos and Hughes' investigation, this study, by using a multi-method approach of data collection, reveals further sources of tutors' confidence or lack of confidence and the relation between confidence, cognition, and action.

5.4.3 Trust and the membership within the community of practice

It seemed that trust was another source of confidence, in that it was a confirmation of tutors' qualification for the job. Most of the tutors in this study were master's or doctoral students who had demonstrated their competence in their subject area, thus earning the trust of their lecturers, which is reflected in their choice of these students as tutors on their courses. However, most tutors were still students who were not formally employed by the university, which placed them on the periphery of the community of practice. Tutors were aware that students might distrust their assessment when the result of assessment did not meet students' expectations, especially when students received more positive comments on the written scripts from a different tutor or a lecturer. Distrust may also result in worry and lack of confidence. This finding concurs with the findings of Carless' (2009) study that distrust increases the difficulty of keeping professional standards of assessment, especially when there is a possibility of failing students. In addition, Lee and Schallert (2008) note that trust between teachers and students affects the effectiveness of feedback. Therefore, distrust may partly account for the tutors' uncertainty of the effectiveness of feedback.

According to Lave and Wenger (1991), novices in a community of practice usually carry out less difficult tasks; therefore, their participation is a process of increasingly legitimate peripheral participation. However, in the present study, the tutors carried out the same activity of assessment as those who supervised them. The tutors received instructions on how to carry out the activity but they had no professional training on assessing writing, therefore, the tutors were likely to face distrust at work and worry about their performance. The contradictions within the community of practice and the need for professional support for the tutors will be further discussed in the following section.

5.4.4 Summary

This section analysed the ECA interactions at the collective level of the activity of assessment. It pointed out that emotions were distributed between the tutors and the students via feedback. On one hand, the tutors' assessment feedback influenced students' subsequent emotional and cognitive reactions; on the other

hand, the students' evaluation and response to the tutors' work influenced the tutors' emotions at work. Empathy towards students was the most common emotion among the tutors. Empathy derived from the tutors' memory of previous emotional reactions to the feedback on their own work, which affected their current practice. The tutors intended to distribute positive emotions among students via their assessment feedback. Another emotion, confidence, not only related to the tutors' working experiences but also related to their status in the community of practice and with the students, as well as those who supervised their work. Tutors' confidence may also result from the trust from those who supervised their work and from their students. Distrust also related to the tutors' marginal status in the community of practice and was a barrier to the tutors' work. The next section will analyse the contradictions of the activity system that need to be reduced in order to better regulate ECA in the activity of assessment.

5.5 Regulation of ECA and contradictions in the activity system

The present study found that the tutors had convergent and divergent beliefs and practices, the extent of which resulted in emotional reactions. The convergent beliefs and practices were mainly derived from the rules and tools of assessment distributed among the tutors and the interactions between individual tutors and other members of the community of practice within the current activity system. The distribution of rules and tools of assessment aimed to regulate the tutors' practices to achieve consistency in assessment.

However, the analysis of the interaction between ECA reveals that the distribution of cognition was not efficient in regulating the following aspects. Firstly, the new tutors were not confident in carrying out the assessment activity: they had difficulties in deciding on a fail grade and/or providing feedback on writing. Secondly, divergent beliefs and practices remained among the tutors regarding commenting on and assessing formal aspects of written work, and the construction of positive feedback. These divergences derived partly from the tutors' previous learning and teaching experiences and partly from the uneven distribution of cognition among tutors via individual interactions within the current activity and the tutors' interactions with others beyond the current activity system. Finally, divergences also exist between beliefs and practices, especially between the believed goal of helping students making improvements and, in practice,

166

justifying a grade via written feedback. These issues indicate that there is a need to improve the quality and flow of information of the current activity system. According to Engeström (1987, 2001), the improvement of an activity can be achieved via expansive learning. The starting point of expansive learning is the analysis of the contradictions within the activity system.

The divergence between beliefs and practice among the tutors will be discussed in relation to the regulation and contradictions in the activity system in this section in the following sequence: 5.5.1 Regulation and contradiction in the top triangle of the activity; 5.5.2 Regulation and contradiction in the expanded model of the activity; 5.5.3 Summary.

5.5.1 Regulation and contradictions in the top triangle of the activity

The main contradiction within individual tutors was the believed goal of feedback as helping students to improve and the actual, practised goal of justifying the grade awarded. In addition, the tutors also had divergent beliefs and practice regarding error correction. These divergences can be analysed with the help of the top triangle of activity theory developed by Leont'ev (1978) (Figure 2.1).
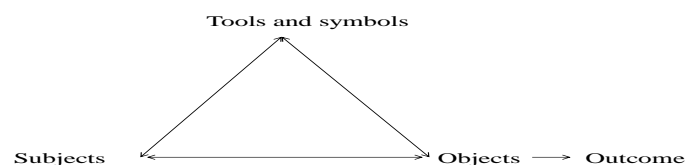
Figure 2.1: Original model of activity theory (Leont'ev, 1978)

According to this model, the individual tutors are subjects of the activity. The object is the written work by undergraduate students. The tools of assessment include criteria, marking schedules, guidelines of marking, and pencils or pens. Language is also a tool, by which the tutors composed assessment feedback.

The tutors in this study believed that the intended goal of providing feedback was to help students improve their written work next time, although they were not sure whether such improvement would happen. The tutors tended to indicate most, if

167

not all, grammatical errors in students' assignments. This practice was in contradiction with their beliefs about students' possible negative emotional reactions to error correction but was in agreement with their belief that error correction may help students to improve. However, the tutors usually provided error corrections randomly across the written work without clarifying the sources of errors, providing instructions on how to avoid certain types of errors, or suggesting any specific goals for individual students to achieve regarding the formal issues of writing. It seems to be a common issue in the present study, and in others, that advice to students on how to improve requires more professional knowledge and effort than merely providing praise and unsystematic corrective feedback. What is more, the tutors' contradictory belief and practice in regard to error correction was embedded in the criteria they used: although grammatical errors or quality of expression were among the assessment criteria, there were usually no implicit statements on whether to correct, how much to correct, or how to correct such errors effectively.

According to Hattie and Timperley (2007), feedback should include information on how to improve. However, the present study found that there was little evidence in these tutors' feedback on how the students' written work might be improved. This finding is similar to Orsmond and Merry's (2011) finding that some tutors in British universities tended to use praise and corrective feedback but provided no information on how to improve. The overall comments provided by the tutors were generally composed of positive and negative points according to the criteria, or simply praise for good written work, or no overall comments at all. However, as indicated above, it has been found in the literature that encouraging words do not in themselves cause improvement (Leki, 1990), although they may have a positive effect on confidence (Beason, 1993; McGee, 1999). It seemed impossible for students to improve when the encouraging words were simply slogans such as "Well done" as was provided by some tutors in the study. Strategies suggested in the literature to improve the effectiveness of feedback include: integrative feedback (e.g. Broad, 2003), assessment dialogues (Carless, 2006; Nicol, 2010), portfolio assessment (e.g. Weigle, 2007), dynamic assessment (Poehner & Lantolf, 2010), and progressive feedback on drafts (Duijinhouwer et al., 2010); such strategies, however, were not evident in this study as the one-way feedback was only given on the final submission of the written work.

On the other hand, efforts and strategies for keeping consistency of assessment were taken seriously by the tutors in this study to achieve the summative function of assessment. To make their assessment more reliable, the tutors adopted strategies such as comparing grades of the written scripts of the same cohort of students and adjusted the grades, comparing students' written work before grading and providing overall comments, having several rounds of marking if they had time to do so. There has been argument in the relevant literature that summative assessment aims to provide valid and reliable measurement of learning outcome (Cherry & Meyer, 1993; Moore, O' Neill, & Huot, 2009), that formative assessment aims to engage students with improvement of their work via helpful feedback (Hamp-Lyons, 2008; Sadler, 1998), and that the key to being formative is to provide diagnostic feedback (Hamp-Lyons, 2008). However, some teachers face a dilemma when they are required to achieve both the formative and summative goals (Bailey & Garner, 2010).

In the present study, an analysis of the tutors' feedback revealed that the stated purpose was not achieved in their actual practice as their major concern was to justify the grade and to fulfil the summative purpose of assessment. Thus, this study strongly reinforces the view adopted by others (e.g. Basturken et al., 2004; Borg, 1998; Lee, 2009) that while beliefs may influence practice, they do not determine it. The reason for this divergence between beliefs and practice lies in the double bind (Engeström, 1987, 2001) nature of the goal in the activity system. According to Engeström (1987), double bind is the contradictory information given to the actors. In the present study, the tutors were in a double bind situation in that they were expected to achieve both summative and formative goals in the activity of assessment. Although the tutors were expected to achieve both the summative and formative goal of the activity of assessment, they were not supported by sufficient resources such as time, instructions and training on how these goals could be efficiently and effectively achieved.

The current study concurs with Bailey and Garner's (2010) finding that the tutors were not well informed on how to achieve the formative purpose of assessment. The summative function of these tutors' assessment was dominant. The criteria, marking schedules and the strategies the tutors used were mainly aimed at

fulfilling the summative function of assessment. The assessment feedback could inform students of how well the written work met the criteria or marking schedule, but did not inform students how to improve.

5.5.2 Regulation and contradiction in the expanded model of the activity

Individual tutors were members of a community of practice within the broader activity system of the university. The divergent beliefs and practices which existed among tutors reflected the contradictions between the beliefs and practices within the community of practice. The present study found various contradictions within the community of practice among tutors, and between tutors and their seniors. For example, in the open-ended questions in the survey, one tutor wrote that the lecturer of one course s/he was tutoring was helpful whereas the lecturer for another course was not. Interview data also demonstrated that tutors were aware of the different practices between lecturers and between tutors. For example, Mia found that she gave "slightly higher marks" (Mia, interview) than the other two tutors of the same class. George believed that tutors "tend to be slightly easy" (George, Interview) if they knew that a student was working hard. Data from focus group sessions also demonstrated that tutors sometimes received contradictory instructions on their work from different lecturers. These contradictions can be analysed according to Engeström's (1987) expanded model of activity theory.

One reason for the contradictions is the division of labour in the community of practice. As can be illustrated by the following triangle (Figure 5.3):
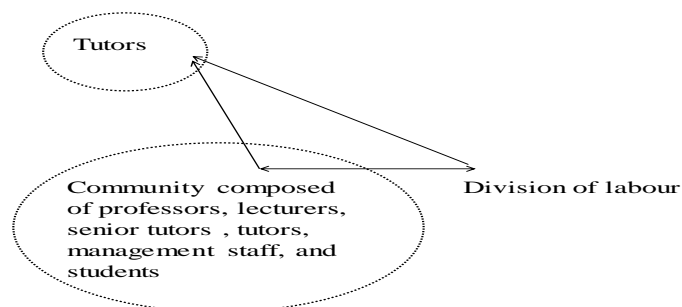

Figure 5.3: Division of labour in the activity of assessment

Firstly, there were divergent divisions of labour among tutors. In the present study, some tutors were also responsible to assist lab sessions or give tutorials, while other tutors only assessed students' written work without any direct contact with students. Some tutors were also paid to sit in the classes in order to be familiar with the requirements and content of the courses, but most tutors did not attend classes during the period of study, which, according to one tutor, was because they were not paid to. The differences in the division of labour resulted in divergent interaction between tutors, students, and the course work to be assessed. Secondly, there were divergences in the division of labour between tutors and their supervisors. Tutors were supervised by senior tutors or lecturers. The tutors and lecturers both assessed students' work. However, lecturers, unlike the tutors, were fully involved in other pedagogic activities such as course design and teaching. This difference was likely to make tutors less authoritative in assessment as they were not as familiar as their lecturers with the students, the course content, and the requirements or criteria of assignments. As one tutor pointed out, it was the lecturers' course and the role of the tutors was to assist the lecturers' goals; whether or not they agreed with the lecturers did not matter. In addition, most tutor participants were students themselves. They assessed students' assignments within the same broad time-frame as their own assignments were assessed by their lecturers. They had concurrent experiences of assessment from opposite positions due to their dual identity. Their identity as students made them less authoritative as lecturers and more empathetic towards students and sometimes more lenient in grading. Thirdly, there were divergences regarding whether or how much errors should corrected. Some tutors mentioned that they would suggest to students whose grammatical competence was poor that they should see tutors in the learning support unit in the university, implying that they did not consider that it was their responsibility to do this. Error correction seemed to be evidence of extra effort tutors took at work. However, how many errors to correct usually depended on how much time tutors could afford and how competent they felt they were to correct errors.

The second reason was the inappropriate rules and the ineffective distribution of the rules, as illustrated in Figure 5.4:
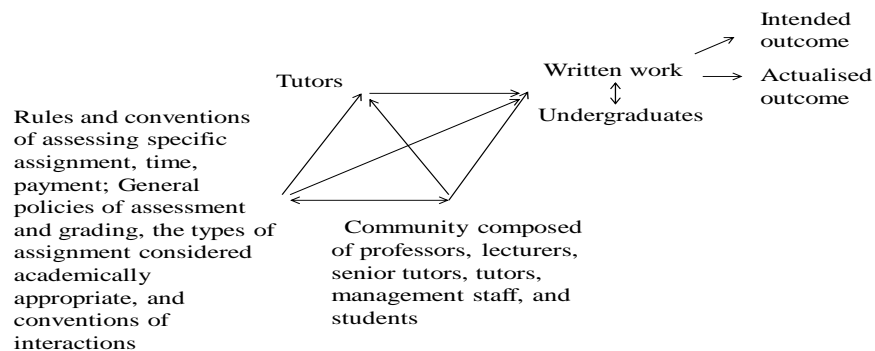
171

Intended
outcome

Written work

Actualised
outcome

Tutors

Rules and conventions
of assessing specific
assignment, time,
payment; General
policies of assessment
and grading, the types of
assignment considered
academically
appropriate, and
conventions of
interactions

Undergraduates

Community composed
of professors, lecturers,
senior tutors, tutors,
management staff, and
students

Figure 5.4: Rules for the tutors to carry out the activity of assessment

The divergence of beliefs among the tutors regarding some practical issues such as error correction, weighting grades, and providing feedback positively indicated that there was no clear policy of, or instruction about, such issues. The conventions regarding the practical issues were usually distributed via oral interactions between tutors and lecturers. For example, one tutor was advised by the supervisor that the length of an overall comment on an essay should be, on average, four lines, while other tutors were not similarly advised by their supervisors. The interactions augmented tutors' individual cognition and were helpful for tutors to solve practical issues emerging at work. However, these rules had a limited scope of distribution (usually between those tutors who interacted frequently with each other) and were not distributed evenly across the wider community of practice. The analysis points to the need for rules and conventions (as manifest in holistic criteria and discussions between lecturers and tutors) to be sufficiently clear and practical if the activity of feedback is to be effective.

The rules of assessment also included payment for the tutors for the time they spent on assessing each piece of written work. The payment for tutors was considerably lower than lecturers. The average payment was $16 per hour and the paid time for assessing an essay was about twenty minutes. Moreover, they were only temporarily employed on a part-time basis for the duration of a semester; they had no guarantee of work beyond that immediate horizon. This was in stark contrast to the (relatively) well paid and secure positions held by their seniors. The division of labour outlined above demonstrates the hierarchical structure of

172

identity, payment, access to knowledge and contact with students within their community of practice, although the tutors carried out the same activity of assessment as the lecturers. The tutors, who were novices in the community of practice, often needed more time than they were paid to fulfil their work. This limitation of rules contributed to the difficulty for them in achieving the double bind goal of the activity.

Finally, contradictions may also result from the inefficient information flow within the current academic community, as suggested in Figure 5.5:
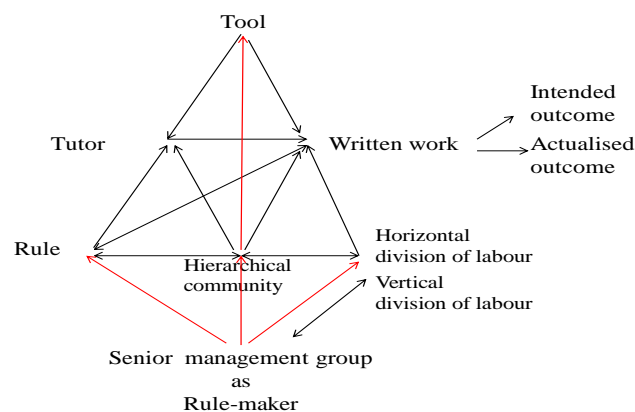
Figure 5.5: Information flow in the current community of the assessment activity

As depicted in Figure 5.5, the main direction of cognitive distribution within the community was from the centre to the periphery due to the predominant vertical division of labour. To some extent, there was a horizontal division of labour in the activity in that tutors shared some of the work of assessing undergraduate assignments with their lecturers as well as fellow tutors. However, these tutors were at the bottom of the vertical distribution of labour and power because they were on the periphery of the community of practice. For example, they were expected to follow the criteria, marking schedules, guidelines and rules of assessment from their supervisors who also provided oral instructions. These rules were governed by the overall university's assessment policy decided by a sub-group of the Senior Management Group who were rule-makers of the current activity system.

173

However, this one dimensional information flow could not cater to the fact that, in many cases, the tutors were executive decision-makers in carrying out the community's rules. Their decisions were reached according to their individual histories. Each tutor had unique cognitive and emotional experiences of receiving assessment feedback on their own written work by their teachers. Moreover, the tutors had different academic backgrounds: some tutors had specialized in English or other language-rich subjects while other tutors had not. The tutors were also at different stages of cognitive development in their disciplinary studies and had different tutoring experiences. Therefore, these individual differences gave rise to divergences in beliefs, and when these were put into practice they led to contradictions within the community of practice (Lave & Wenger, 1991)

The contradictory practices within the community may cause further contradictions between tutors and students. Focus group data demonstrated that sometimes the tutors were in an embarrassing situation because the lecturers did not give them the same amount of information they gave to students. Moreover, contradictions occurred when students received different evaluations on the same piece or different pieces of assignment from different tutors or different opinions between a tutor and a lecturer. For example, Mark had experience of receiving opposite opinions on the quality of expression in two pieces of his own written work submitted during the same period; Henry had a student who asked him for further explanations of the assessment because a lecturer offered a different opinion. The tutors' individual beliefs, although influencing their practice, were usually not able to be articulated, discussed, or distributed formally in the community. As one tutor commented, her opinion did not count. The information flow that was crucial for the achievement of the intended goal of the assessment activity yet was missing in the current activity system is demonstrated in Figure 5.6:
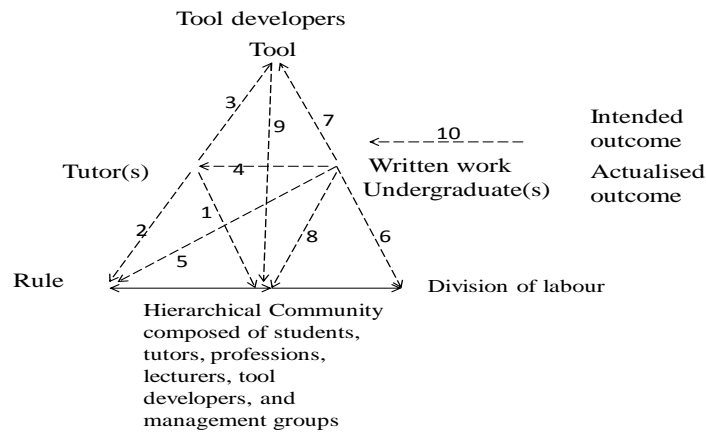
Figure 5.6: Information flow needed in the current community

Figure 5.6 demonstrates that there was a need for information from the tutors, the students, and tool developers to all members in the current community. Firstly, there should be more channels or opportunities for the tutors to inform the academic community (arrow 1). The tutors should be included in the process of rule-making (arrow 2). They should also be able to participate in the tool developing activity and inform the relevant research (arrow 3). In addition, students' feedback on the tutors' practices should not only go to the management groups by filling in a summative appraisal form at the end of the course: there is a need for the tutors to receive formative feedback from the students throughout the course (arrow 4) so that the tutors could be better informed of, and cater to, the students' needs. Students should also be able to inform the rule-making process (arrow 5), the division of labour (arrow 6), the tool development (arrow 7), and had their voice heard in the community (arrow 8). Researchers, as tool developers, should also be integrated in the community (arrow 9). Finally, the outcome of the activity should in turn inform the members of the community for future improvement. However, in the current activity system, there was not such evidence that the rules of assessment were based on the co-construction by all the supposed members of the community including tutors, students, and researchers.

These contradictions point to the need of a new model of interactions within the community (Figure 5.7).
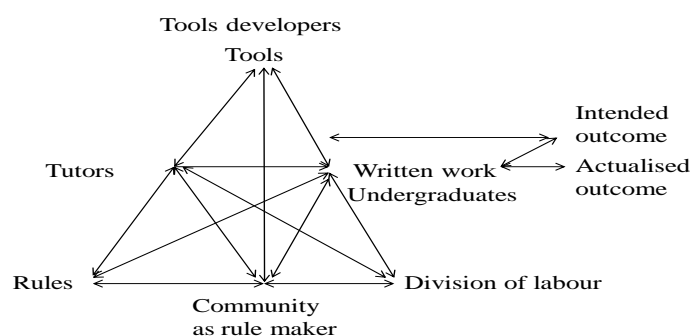
Figure 5.7: Model of information flow in the assessment activity

In the suggested model of activity, the information flow is bi-directional between the components of the activity system. The information within the activity system is transparent to all members of the community who actively co-construct the activity by providing feedback to each other and negotiate goals, rules, and appropriate tools in the multi-voiced community in order to achieve the intended outcome. In turn, the intended and the actualized outcome informs all the members of the activity for their future improvement.

The analysis above reveals how division of labour, rules, and the interactions within the community of practice led to divergent beliefs and practices of these tutors. Among the limited number of studies into university tutors' divergent and convergent beliefs and practices of assessment, Tang and Harrison (2010) found tutors in an English programme had convergent beliefs that corrective feedback was effective but were not sure if students would use the feedback they provided. The divergent beliefs among tutors in Tang and Harrison's (2010) study were the formative and summative roles of assessment and feedback, which dilemma also confused tutors in the present study. However, unlike what was found by Tang and Harrison, tutors in this study had convergent beliefs about the importance of the formative role of feedback, but demonstrated in practice their primary concern was to justify the grades they allocated. The difference between the findings of these studies perhaps is partly because Tang and Harrison only explored beliefs by survey and interview without collecting data of actual practice, and they did not provide explanations for the divergences and convergences they uncovered. Orrell

(2006) examined the assessment practices of experienced teachers. The convergent practice among the teachers in that study was that feedback did not suggest further revision, which was also the case with tutors in this study, as was the divergent length of feedback provided. However, Orrell concluded that the divergence between teacher practices in his study was because of their different assumptions of their identity in the assessment process. In some other studies of teacher cognition of assessment, it was found that institutional policies and situations and modular patterns of course delivery constrained teachers' practices (Bailey & Garner, 2010; Price et al., 2010; Meyer et al., 2010). Other reasons - such as raters' various backgrounds (Hamp-Lyons, 1991; Sakyi, 2000; Vaughan, 1991) or criteria difference (Barkaoui, 2007) - have also been found as causes of divergent practices of assessment.

However, none of these studies has reported an in-depth analysis of the sources of contradictions between beliefs and practices within a community of assessment practice. These sources include the various cognitive backgrounds and identities of the members of the community, the power relationships and interactions between them, the rules, tools, and goals of the activity. It is necessary to take a holistic view of the dynamic interaction of these factors to explain the convergences and divergences of the beliefs and practice within a community of practice. Thus, the analysis of the findings of the present study contributes to a greater understanding of assessment activity. In addition, the analysis of distributed cognition within a community of practice contributes to teacher cognition studies about the sources of cognition. Previous studies have found that the main sources of university teachers' beliefs are their former experience as learners and later teaching practice (e.g. Boice, 1992; Dunkin, 2002). This study confirms the previous findings but also reveals more sources, the most important of which is the distributed cognition among the members of the community of practice.

The contradictions within the community were embedded in the wider context of the activity in the university. The tutors needed professional support that could help them to better regulate their cognition and emotions to achieve the goals of assessment. This requires negotiations between the current assessment activity and the activities of policy-making, professional development, and the

177

development of contextualized theories of assessment. This will be discussed in the next chapter.

5.5.3 Summary

This section discussed the divergences between tutor beliefs and practices of assessment from the perspectives of social regulation (Vygotsky, 1981) and contradiction (Engeström, 1987, 2001). The major divergences and convergences between tutors' beliefs and practices and among tutors were analysed firstly in relation to the top triangle of activity developed by Leont'ev (1978), and then according to Engeström' s (1987, 2001) expanded model. It argued that the divergences resulted from the double bind goal of assessment, the implicit or inappropriate rules, the ineffective tools, and the insufficient information flow in the community.

5. 6 Conclusion of discussion

The findings of the study have been discussed mainly from the perspective of the expanded notion of regulation and mediation (Vygotsky, 1978) and activity theory (Engeström, 1987). It has been argued that tutors' cognitive development is contextually regulated and emotionally modified. This conclusion is composed of the following points. First, by applying Engeström's (1987) activity theory to the study, the context of tutor cognition of assessment is defined as the activity system which is influenced by the policy-making activities and tutor development activities and evaluated against the theoretical models of the activity. The context was the source of both cognition and emotion. The emergent ECA (emotion, cognition, and action) interactions in the context were analysed at the levels of operations and actions in the activity according to the think-aloud and stimulated recall. It was argued that tutors regulated their emotions in order to achieve the goal of assessment. Tutors were also aware of students' emotional reactions to the assessment and feedback; therefore they tried to regulate students' emotional reactions by positive feedback. This argument expands Vygotsky's theory of other and self- regulation of cognition to the more complex level by explaining the regulation of emotion. The historically developed ECA interactions in the community were then analysed. Finally, it was argued that the contradiction within the activity system was the major source of cognitive dissonance

(Festinger, 1957) and negative emotions of individuals, which was a barrier to achieving the intended goals of assessment.

This present study contributes to academic understandings in three ways: theoretical, methodological, and practical. Firstly, it contributes to the theoretical development of teacher cognition by applying a holistic Vygotskian activity theory perspective to the analysis of the roles that context play in teacher cognition and how individual cognition interacts with the contextual factors. It also expands activity theory in that it tries to explain how individuals' ECA are regulated by self and others. Secondly, it has a methodological contribution to the study of teacher cognition, especially regarding the combination of think-aloud and stimulated recall in the exploration of ECA interactions in the process of assessing writing in the real context. Thirdly, it has practical implications for the activity of assessing disciplinary writing, particularly in regard to the understanding of the tutors' beliefs and practice and the contradictions within the context and the support that is needed.

The next chapter will conclude the thesis by discussing the theoretical, methodological, and practical implications of the study.

CHAPTER SIX: CONCLUSION

6.0 Introduction

This chapter concludes the thesis. It will firstly overview the scope of the study including the purpose of the study, the process of data collection and analysis the findings, the discussion of the findings and the limitations (Section 6.1). It will then discuss the implications of the findings in terms of tutor development, policy-making, research, and theoretical development about tutor cognition of the assessment activity (Section 6.2). Finally, it will briefly present my reflections on conducting this study (Section 6.3).

6.1 The scope of the present study

As a case study about the subject tutors in a single faculty in a New Zealand university in a particular period of time, the present study did not aim to make generalisations, except that the findings may be relevant to future studies on tutor or teacher cognition in similar contexts. The purpose of the present study was to explore tutors' beliefs about and practices in assessing undergraduates' written work, and to critically examine the crucial factors that influenced these.

To answer the research questions, a multi-method approach was used to collect data from October 2008 to October 2009. Firstly, a preliminary online survey was used to elicit tutors' attitudes toward feedback and to recruit participants. Secondly, sixteen tutors who volunteered to participate in the study were interviewed regarding their routines of providing assessment feedback, their beliefs of good practices, and the sources of their beliefs. Thirdly, nine of the interviewed tutors participated in the think-aloud and stimulated recall sessions: each participant verbalised their thoughts aloud while marking one or two of their students' written work in the presence of the researcher; then the tutors discussed with the researcher the issues that affect their decision-making in the think aloud sessions. Finally, two focus group discussions were conducted with seven of the interviewed tutors, four of whom also were also participants of the think-aloud and stimulated recall sessions. The topics of focus group discussion were the topics emerged from data analysis of survey, interview, think-aloud and stimulated recall sessions. The audio -recorded data were transcribed by myself.

The transcripts were checked by a fellow researcher. The data were analyzed by a grounded-theory approach, firstly with the use of NVivo8 and then checked manually by myself and discussed with my supervisors.

One of the key findings of study was that the tutors believed their purpose of providing feedback was to help students to make improvement; however, the purpose their actual practice served was to justify the allocation of a grade. Another key finding was that the tutors believed that they focused on the content of written scripts yet their attention was attracted by grammatical errors when the quality of expression was low. Moreover, the study found that emotion played an important role in the activity of assessment: on one hand, the tutors' emotions interacted with their beliefs and practices, on the other hand, tutors used feedback to regulate students' emotional reactions.

The findings were discussed from a neo-Vygotskian socio-historical activity theory perspective. It was argued that tutor cognition was strongly affected by contextual factors in the activity system of assessment. The emotions interacted with cognition and action, and consequently affected the outcome of the assessment activity. The interactions between emotion cognition and action (ECA) is firstly regulated by other people in the activity system and then regulated by tutors themselves. The divergence between tutors' beliefs and practices indicated that other-regulation in the activity system was not sufficient for the tutors to achieve their intended goals. Therefore, there is a need to analyse the contradictions within the activity system and remodel the activity.

One limitation of the study is that the interpretations of the findings were my own, and I am a relatively inexperienced researcher, although, to ensure the quality of my interpretation, the transcripts of the audio-data were checked by a fellow (native English- speaking) researcher in applied linguistics and I was guided by my supervisors. Despite these limitations, the findings of this study suggest a number of interesting implications that will be discussed in the following section.

6.2 The implications of the study

The present study has three significant implications: practical, methodological, and theoretical. The practical implications are the need for professional

development of tutors and the re-examination of assessment policies. The methodological implications relate to the multi-method approach of data collection and the application of the activity theory to the interpretation of findings. The theoretical implications are the analysis of ECA to the development of teacher cognition theories, activity theory, and the theory of assessing writing.

6.2.1 Implications for professional development of tutors

One of the main issues that emerges from this present study of tutors' beliefs and practices is the lack of systematic professional development in the specific area of assessing disciplinary writing, which resulted in their lack of confidence and difficulties in carrying out the activity. As has been discussed in Section 5.5.2, the tutors who were employed to mark assignments were mostly part-time sessional teaching assistants who had knowledge in their own disciplinary area. They had not received formal professional training, nor were they well-informed by relevant literature on assessment or feedback. They were in a marginal place in the community of practice and they were not as fully involved in the teaching and research activities as were the lecturers, although both groups carried out assessment activities. The limited training available to tutors was mainly at procedural level by lecturers or senior tutors. They may have chosen to take some optional workshops or seminars offered by the university's teacher development unit, but no tutor had mentioned this kind of experience. It seemed that any development activity was mainly carried out by the tutors themselves in an ad hoc way rather than as an integrated part of a professional development programme. It can be argued that these tutors were not formal staff of the university and therefore it was not the business of the university to offer them formal training. It is true that such tutors may not take teaching as a career after their graduation because they had their own motives to work as tutors such as gaining work experience or receiving payment. However, the tutors were employed to promote important goals of the activity system. The formative pedagogical goal of assessment, as has been discussed in Section 5.3.1 and Section 5.3.2, was essentially unachieved. In regard to the summative goal of assessment, although criteria were used in an attempt to maintain consistency, examples of inconsistent assessment results were found in the community of practice. One reason for this is that training in the use of criteria is a necessary condition to maintain consistency

(Reddy & Andrade, 2010), and this was not formally provided for the tutors. If the university could not ensure systematic training for tutors, there could be no guarantee that assessment activities carried out by tutors would maintain institutional or professional standards. On the other hand, tutors who had been working for years did gain very valuable contextual knowledge and had developed some techniques of assessment. They also had accumulated knowledge and beliefs of assessment and feedback in their own practice. However, they were not sure whether their beliefs and practices were professionally sound because they were at the margin of the community where their professional opinions were hardly ever solicited. Neither were they involved in any relevant research to develop more rational theories of their practice, nor were their knowledge and experiences explicitly valued in the development of theories or policies.

The need for professional development found in this study concurs with the findings of DeLuca and Klinger (2010) and Volante and Fazio (2007) that there is a lack of assessment literacy among university teachers in general. Morevoer, it supports the findings made by Sutherland (2009) that tutors not only made contributions to their teaching communities because of their dual identity (as they were simultaneously students) but also benefited from this experience of teaching. However, the tutors did not receive any systematic professional training, which is the main reason of their non-authoritative identity and marginalized situation in teaching community.

The present study concurs with previous studies that systematic professional training is urgently in need of ensuring the quality of activities such as assessment. Furthermore, by revealing the practical difficulties that the tutors met in assessment, the present study provides detailed information about the areas on which the professional development should focus for tutors to carry out assessment activity. According to Engeström (1987, 1999b, 2001), expansive learning includes questioning, analysing the current contradictions in the activity system, and remodelling the activity with new goals to achieve. The establishment of a new model requires a collaborative effort of people who are involved in the activity of assessment. Arising from the present study is a proposed onsite professional development model for tutors in the present activity system, which is

based on the findings of the present study and inspired by Engeström (1987, 1999b, 2001) concept of expansive learning.

This model would include expertise and collaborative efforts of those who work in subject areas as well as in education and applied linguistics. The goal of development would be negotiated with the relevant tutors, catering for their specific needs and integrating their academic and career goals with the goals of the assessment activity. Educational experts would focus on pedagogy and strategies of assessment. These experts may also provide strategies on the regulation of emotions in the assessment activity. Applied linguists would provide specific knowledge on language, writing, and feedback. Subject experts would share their knowledge on subject content course related information. The professional development needs to be systematically designed and make access to tutors who can participate at a regular basis with specific goals to achieve at different stage of development.

Meanwhile, the tutors should have ready access to the experts for consultation of issues that emerge at work. In addition, tutors would be encouraged to share their beliefs about and practices in the assessment activity so that they would co-construct solutions to contextual problems they encounter, and distribute their knowledge and insights among their peers. The tutors would also have structured opportunities for dialogues with students, in which they would apply their developed knowledge to the on-going academic relationships, and receive feedback from students to inform their further practice. Finally, the tutors would have guided reflection on their practice to theorise their personal knowledge and opportunities to disseminate their theories to the wider academic and professional communities. These efforts of professional development aim to fully integrate tutors in the academic community of practice in the assessment activity and to ensure the quality of their professional practice. The tutors should be informed of offsite possibilities and onsite opportunities of professional training. This systematic professional development should be part of institutional support provided for the tutors and teachers, and such provision requires the re-examination of institutional policies and resources, which will be discussed in the following section.

6.2.2 Implications for institutional policy

The present study found that the tutors' emotion, cognition, and action were regulated by the contextual factors of the activity, which in turn were determined by institutional policies. The activity of assessment carried out in the university was based on the policies of assessment distributed across the university, which focused on procedural aspects of assessment. These procedural policies could be found on the website of the university, and in various forms of staff manuals and marking requirements in different departments. These central polices had priority against local policies, and determined the summative role of assessment. Tutors not only followed these policies but were also aware of the institutional consequences of their assessment. In focus group discussions, the tutors pointed out that these assessment policies influenced both students' evaluation of the course (which was often carried out in the form of questionnaire from the teacher development unit at the end of a course) and the enrolment of students in each department. Therefore, tutors tended to be lenient when deciding grades. However, they were not sure whether the organisational needs matched the pedagogic goals of assessment. The divergence between tutors' beliefs and practices concurs with the findings in the existing literature that the dominant goal of assessment is measurement (Bailey & Garner, 2010; Meryer, et al., 2010; Price et al., 2010), that teachers are under institutional and program constraints (Bailey & Garner, 2010; Price et al., 2010) to achieve the summative goal of assessment.

The institutional policies also affected the application of alternative models of assessment. In the reviewed literature, various models and strategies were suggested as more effective to achieve the formative goal of assessment. These models included integrative feedback (e.g. Broad, 2003), portfolio assessment (e.g. Weigle, 2007), and assessment dialogues (Carless, 2006). However, none of the approaches were used by the tutors in the current activity of assessment. The reason is that the application of these models requires not only knowledge of assessment but also negotiation between policy-makers and practitioners about the activity of assessment. In the context of the present study, it appeared that the types of assignment and feedback, the rules and tools of assessment, the terms and conditions of staff employment, division of labour, and appraisal of teachers were all decided unilaterally by the policy-makers in the university. According to Engeström (1987, 2001), the application of new models of assessment requires

co-operation between people participating in relevant activities to reduce the quaternary level of contradictions (Engeström, 1987, 2001). Therefore, institutional policies of assessment should be re-examined and adjusted to address these issues.

One approach to identifying policies that can better address the practical issues is to integrate policy makers into the professional development programme, from which the policy makers can receive fresh information from multi-voices of practitioners, experts, and students. Moreover, opportunities and responsibilities should be given to all members of the community to advise policy makers about the effects of the current policy in local contexts, the issues to be solved, and the expected outcomes.

6.2.3 Implications for research on teacher cognition and assessment activity

The implications discussed in Section 6.2.1 and Section 6.2.2 required changes to or developments in the activity system as necessary processes of expansive learning. To properly prepare for such changes, in-depth study is needed to fully explore the complexity of the context, monitor the changes, and provide research information to facilitate the process of expansive learning.

To achieve this purpose, multi-methods of data collection are needed to explore different aspects of the activity. Conventional studies on teacher cognition and assessment activity largely depend on survey and interview data, and self-report data may not be reliable, or even honest. The present study tried to solve this issue by applying an in-depth multi-method approach of data collection, which included an on-line survey, interview, think-aloud, stimulated recall, and focus group discussions. The purpose of using these methods was to collect ongoing data of both beliefs and practices of tutors as members of a community of practice. The emergent data were constantly compared and contrasted with each other in a triangulated process of grounded analysis, which provided the opportunity to explore tutor cognition and practice in depth. A suggestion for future studies is that the various methods of data collection need to be carefully selected to triangulate and compensate each other to holistically serve the research purposes.

186

Another suggestion is that, in order to capture thought in action – 'cognition in flight', as Vygotsky termed it - further studies need to explore the use of introspective methods such as think aloud in real contexts of practice, rather than in experimental or quasi-experimental settings.

Regarding the important role that context plays in teacher cognition, further studies need to adopt a systematic framework that can reveal the complexity of any institutional context.

6.2.4 Implications for the theoretical development of teacher cognition, activity theory and assessing disciplinary writing

The present study has several theoretical implications. The most significant implication of the study is that it reveals the interactions between emotion, cognition, and action in the activity of assessing undergraduate's written work. This analysis contributes to the development of teacher cognition theories, activity theory, and the theory of assessing disciplinary writing.

Previous studies about teacher cognition focus on the relation between beliefs and practices (Clark & Peterson, 1986; Borg, 2006), and the role of emotion in assessment activity has been explored only by a very limited number of studies which have focused on specific emotions such as empathy (Värlander, 2008), trust (Carless, 2009; Lee & Schallert, 2008), and confidence (Goos & Hughes, 2010), and emotional reactions in the process of marking (Crisp, 2008). However, none of the studies have explored the interaction between emotion, cognition, and action in the activity of assessment in an in-depth and systematic way. The present study found that emotions such as empathy, confidence, trust, frustration, worry, and other emotional reactions were derived from the interactions between individual cognition and contextual factors. These emotions interacted with tutors' cognition and manifestly affected their decision making in assessment. In addition, the tutors tried to regulated students' emotions by providing positive assessment feedback. By explaining the mechanism of the interactions between ECA (emotion, cognition, and action) from the perspective of activity theory, the present study contributes to the theoretical development of teacher cognition studies. Future studies need to further explore the ECA interaction in pedagogic

activities such as assessment in other contexts. A broader theoretical construct of cognition is needed which incorporates the important role that emotion plays in thinking and decision-making. In addition, the current expanded model of activity theory does not explicitly depict the role emotion plays within an activity system. The explanation of ECA interaction in this study contributes to the development of activity theory. Finally, the analysis of ECA interaction has implications to the theoretical development of assessing writing. Conventional studies on assessment have focused on the objectivity of assessment. Emotion in assessment is often neglected or regarded as an issue to be avoided. However, the present study reveals that emotion played an important role in assessment activity: on one hand, it affected tutors' decision-making and the outcome of assessment; while on the other hand, it is used by the tutors to regulate students' ECA. Therefore, more studies are needed to explore the important roles that emotion plays in assessment, and to reveal how emotion can be better self- or Other-regulated to achieve both the formative and summative goals of assessment.

In addition to the ECA interactions, the present study contributes to the literature regarding the gaps in research into providing assessment feedback for disciplinary writing for undergraduates, and the effort to draw together the discussions in different discourse communities. Further studies are needed to focus on the coherence of the discourses and to address the cross- disciplinary nature of providing assessment feedback.

6.3 Final reflection

By exploring tutor cognition of assessing disciplinary writing, I have made the following personal discoveries: Firstly, I have come to understand that there is natural connection and coherence between emotion, cognition, and action, which cannot be neglected in the study of beliefs and practices. This understanding provides insights for my own professional and academic development as well as my emotional well-being at work. Secondly, there is connection and coherence between different branches of studies on tutor cognition of assessing writing, and my starting point from applied linguistics has been greatly expanded by considering perspectives from related disciplines, and I now realise the need to integrate the branches of studies. By applying, for the first time the framework of activity theory, I have been able to draw connections between these branches.

Thirdly, to embrace difference branches of studies and to explain the complexity of the ECA interaction, a systematical analytical framework, such as the Vygotskian activity theory approach is needed. Finally, I have learnt that a multi-methods approach of data collection and analysis is required to catch the complex composition of different aspects of professional activity, and feel that I have developed my research skills. By doing this study, I have indeed travelled far on my academic journey but I realise that I still have a long way to go.

REFERENCES

Adler-Kassner, L., & Harrington, S. (2010). Responsibility and composition's future in the twenty-first century: Reframing accountability. *College Composition and Communication*, *62*(1), 73-100.

Bailey, R., & Garner, M. (2010). Is the feedback in higher education assessment worth the paper it is written on? Teachers' reflections on their practices. *Teaching in Higher Education*, *15*(2), 187-198.

Baker, B. A. (2010). Playing with the stakes: A consideration of an aspect of the social context of a gatekeeping writing assessment. *Assessing Writing*, *15*(3), 133-153.

Barkaoui, K. (2007). Rating scale impact on EFL essay marking: A mixed-method study. *Assessing Writing*, *12*, 86-107.

Barlow, L., Liparulo, S. P., & Reynolds, D. W. (2007). Keeping assessment local: The case for accountability through formative assessment. *Assessing Writing*, *12*, 44–59.

Barnard, R. (2010). Activity theory: A framework for analysing intercultural academic activity. *An International Journal of Human Activity Theory*, *3*, 1-13.

Bassey, M. (1981).Pedagogic research: On the relative merits of search for generalization and study of single events", *Oxford Review Of Education*, *7*(1). 73-93.

Basturkmen, H., Loewen, S., & Ellis, R. (2004). Teachers' stated beliefs about incidental focus on form and their classroom practices. *Applied Linguistics*, *25*(2), 243-272.

Bazeley, P. (2007). *Qualitative data analysis with NVivo*. London, UK: Sage.

Beason, L. (1993). Feedback and revision in writing across the curriculum. *Research in the Teaching of English*, *27*(4), 395-422.

Becker, H. J. (1991). When powerful tools meet conventional beliefs and institutional constraints. *The Computing Teacher*, *18*(8), 6-9.

Bell, J. (1999). *Doing your research project: A guide for first-time researchers in education and social science*. Buckingham, UK: Open University Press.

Berscheid, E. (1987). Emotion and interpersonal communication. In M. E. Roloff & G. R. Miller (Eds), *Interpersonal processes: New directions in communication research* (pp. 77-88). Newbury Park, CA: Sage.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability*, *21*, 5–31.

Biggs, J., & Tang, C. (2007). *Teaching for quality learning at university* (3rd ed.). Maidenhead: Open University Press/McGraw Hill.

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, *17*, 102–118.

Bogdan, R. C. & Biklen, S. K. (1982). *Qualitative research for education: An introduction to theory and methods*. Boston: Allyn and Bacon Inc.

Boice, R. (1992). *The new faculty member*. San Francisco, CA: Jossey-Bass.

Boren, M. T., & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. *IEEE Transactions on Professional Communication, 43*(3), 261-278.

Borg, S. (1998). Teachers' pedagogical systems and grammar teaching: A qualitative study. *TESOL Quarterly*, *32*(1), 9-38.

Borg, S. (2001). The research journal: A tool for promoting and understanding researcher development. *Language Teaching Research*, 5(2), 156-177.

Borg, S. (2006). *Teacher cognition and language education : Research and practice*. London, UK: Continuum.

Borg, S. (2009). Language teacher cognition. In J. C. Richards & A. Burns (Eds), *The Cambridge guide to second language teacher education*. Cambridge, UK: Cambridge University Press.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.

Breen, M. P., Hird, B., Milton, M., Thwaite, A., & Oliver, R. (2001). Making sense of language teaching: Teachers' principles and classroom practices. *Applied Linguistics*, 22(4), 470-501.

Brenner, M. E. (2006). Interviewing in educational research. In J. L. Green, G. Camilli & P. B. Elmore (Eds), *Handbook of complementary methods in education research* (pp. 357-370). Mahwah, N.J.: Lawrence Erlbaum.

Brinkworth, R., McCann, B., Matthews, C., & Nordström, K. (2009). First year expectations and experiences: Students and teacher perspectives. *Higher Education*, 58, 157-173.

Broad, B. (2003). *What we really value: Beyond rubrics in teaching and assessing writing*. Logan, UT: Utah State University Press.

Brown, J. S., Collins, A., & Duguid, P. (1989). Debating the situation. *Educational Researcher*, 18(4), 10-12.

Bryant, A., & Charmaz, K. (2007). Grounded theory research: Methods and practices. In A. Bryant & K. Charmaz (Eds.), *The Sage handbook of grounded theory* (pp. 1-28). Thousand Oaks, CA: Sage.

Burden, J., & Roodt, G. (2007). Grounded theory and its application in a recent study on organisational redesign: Some reflections and guidelines. *SA Journal of Human Resource Management*, 5(3), 11-18.

Burns, A. (1996). Starting all over again. In D. F. J. Richards (Ed.), *Teacher learning in language teaching* (pp. 154-178). New York, NY: Cambridge University Press.

Burton, V. T. (2010). Activity system, genre, and research on writing across the curriculum. *College Composition and Communication*, 61(3), 584-596.

Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and involvement. *British Journal of Educational Psychology*, 58, 1-14.

Cabaroglu, N., & Roberts, J. (2000). Development in student teachers' pre-existing beliefs during a 1-Year PGCE programme. *System*, 28(3), 387-402.

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education*, 31(2), 219-233.

Carless, D. (2009). Trust, distrust and their impact on assessment reform. *Assessment & Evaluation in Higher Education*, 34(1), 79-89.

Carter, M. (2007). Ways of knowing, doing, and writing in the diciplines. *College Composition and Communication*, 58, 385-418.

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, 12(3), 267–296.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.

Charter, E. (2003). The use of think-aloud methods in qualitative research: An introduction to think-aloud methods. *Brock Education, 12*(2), 68-82.

Cherry, R., & Meyer, P. (1993). Reliability issues in holistic assessment. In M. M. Williamson & B. Huot (Eds), *Validating holistic scoring for writing assessment: Theoretical and emparical foundations* (pp. 109-141). Cresskill, NJ: Hampton.

Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 255–296). New York, NY: Macmillan.

Cohen, A, D & Cavalcanti, M. C. (1990). Feedback on compositions: Teacher and student verbal reports. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 155-177). Cambridge, UK: Cambridge University Press.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education.* (6th ed.). London, UK: Routledge.

Cole, M., & Engeström, Y. (1993). A cultural historical approach to distributed cognitions. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 1-46). New York, NY: Cambridge University Press.

Cooksey, R., Freebody, P., & Wyatt-Smith, C. (2007). Assessment as judgment-in-context: Analysing how teachers evaluate students' writing. *Educational Research and Evaluation*, *13*(5), 401–434.

Cooper, H. M., & Baron, R. M. (1977). Academic expectations and attributed responsibility as predictors of professional teachers' reinforcement behavior. *Journal of Educational Psychology*, *69*, 409-418.

Creswell, J. W. (2007). *Qualitative inquiry & research design: Choosing among five approaches*. Thousand Oaks, CA: Sage.

Crisp, V. (2008). Exploring the nature of examiner thinking during the process of examination marking. *Cambridge Journal of Education*, *38*(2), 247-264.

Crookes, G., & Arakaki, L. (1999). Teaching idea sources and work conditions in an ESL program. *TESOL Journal*, *8*(1), 15-19.

Cross, R. (2010). Language teaching as sociocultural activity: Rethinking language teacher practice. *The Modern Language Journal*, *94*, 434-452.

Crossouard, B. (2009). A sociocultural reflection on formative assessment and collaborative challenges in the states of Jersey. *Research Papers in Education*, *24*(1), 77-93.

Crossouard, B., & Pryor, J. (2008). Becoming a researcher: A sociocultural perspective on assessment, learning and the construction of identity in a professional doctorate. *Pedagogy, Culture and Sociey*, *16*, 221-237.

Daniels, H., & Warmington, P. (2007). Analysing third generation activity systems: Labour-power, subject position and personal transformation. *Journal of Workplace Learning*, *19*(6), 377-391.

Davydov, V. V. (1999). The content and unsolved problems of activity theory. In Y. Engeström, R. Miettinen & R. L. Punamäki (Eds), *Perspectives on activity theory* (pp. 39-52). New York, NY: Cambridge University Press.

DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, *17*(4), 419-438.

Denzin, N. (1984) *The research act*. Englewood Cliffs, NJ: Prentice Hall.

Dheram, P. K. (1995). Feedback as a two-bullock cart: A case study of teaching writing. *ELT Journal*, *49*(2), 160-168.

Diab, R. L. (2005). Teachers' and students' beliefs about responding to ESL writing: A case study. *TESL Canada Journal*, *23*(1), 28-43.

DiPardo, A., & Potter, C. (2003). Beyond cognition: A Vygotskian perspective on emotionality and teachers' professional lives. In V. Ageyev, B. Gindis, A. Kozulin & S. Miller (Eds), *Vygotsky's educational theory in a cultural context* (pp. 317-345). New York, NY: Cambridge University Press.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies.* Oxford, UK: Oxford University Press.

Du Bois, J. W. (2006). *Representing discourse.* Unpublished manuscript, University of California, Santa Barbara. http://www.linguistics.ucsb.edu/projects/transcription/representing.html.

Duff. P. (2008). *Case study research in applied linguistics.* New York, NY: Lawrence Erlbaum Associates.

Duijinhouwer, H., Prins, F. J., & Stokking, K. M. (2010). Progress feedback effects on sudents' writing mastery goal, self-efficacy beliefs, and performance. *Educational Research and Evaluation*, *16*(1), 53-74.

Dunkin, M. J. (2002). Novice and award-winning teachers' concepts and beliefs about teaching in higher education. In N. H. P. Goodyear (Ed.), *Teacher thinking, beliefs and knowledge in higher education* (pp. 41-57). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Ellis, A., & Harper, R. A. (1975). *A new guide to rational living.* North Hollywood, CA: Wilshire Books.

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research.* Helsinki, Sweden: Orienta-Konsultit.

Engeström, Y. (1996). Developmental studies of work as a testbench of activity theory: The case of primary care medical practice. In S. Chaiklin & J. Lave (Eds.), *Understanding practice: Perspectives on activity and context* (pp. 64-103). New York: Cambridge University Press.

Engeström, Y. (1999a). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen & R. L. Punamäki (Eds.), *Perspectives on activity theory* (pp. 19-38). Cambridge, UK: Cambridge University Press.

Engeström, Y. (1999b). Innovative learning in work teams: Analyzing cycles of knowledge creation in practice. In Y. Engeström, R. Miettinen, & R. L.. Punamäki (Eds.), *Perspectives on activity theory* (pp. 377-404). Cambridge, UK: Cambridge University Press.

Engeström, Y., & Miettenin, R. (1999). Introduction. In Y. Engeström, R. Miettinen & R. L. Punamäki (Eds), *Perspectives on activity theory* (pp. 1-18). New York, NY: Cambridge University Press.

Engeström, Y. (2001). Expansive learning at work: Toward an activity-theoretical conceptualization. *Journal of Education and Work*, 14(1), 133-156.

Engeström, Y. (2007). From stabilization knowledge to possibility knowledge in organizational learning. *Management Learning*, *38*(3), 1–5.[not cited in text]

Engeström, Y. (2009). The future of activity theory: A rough draft. In A. Sannino, H. Daniels & K. D. Gutiérrez (Eds), *Learning and expanding with activity theory* (pp. 303-328). New York, NY: Cambridge University Press.

Engeström, Y., & Sannino, A. (2010). Studies of expansive learning: Foundations, findings and future challenges. *Educational Research Review, 5*(1), 1–24.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*, 215-251.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: The MIT Press.

Faigley, L., & Witte, S. (1981). Analysing revision. *College Composition and Communication*, *32*, 400-414.

Falchikov, N., & Boud, D. (2007). Assessment and emotion: The impact of being assessed. In D. Boud & N. Falchikov (Eds), *Rethinking assessment in higher education: Learning for the longer term* (pp.144-156). London, UK: Routledge.

Fang, Z. (1996). A review of research on teacher beliefs and practices. *Educational Research*, *38*, 47–65.

Fathman, A., & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 178-190). New York, NY: Cambridge University Press.

Ferris, D. (1995). Students reaction to teacher response in multiple-draft composition classrooms. *TESOL Quarterly*, *29*, 33-53.

Ferris, D. (1999). The case for grammar correction in L2 writing classes:A response to Truscott (1996). *Journal of Second Language Writing*, *8*(1), 1-11.

Ferris, D. (2003). *Response to students' writing: Implications for second language students*. Mahwah, NJ: Lawrence Erlbaum.

Ferris, D. (2004). The grammar correction debate in L2 writing: Where are we, and where do we go from here? (and what do we do in the mean time...?). *Journal of Second Language Writing*, *13*, 49-62.

Ferris, D. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds), *Feedback in second language writing: Context and issues* (pp. 81-104). New York, NY: Cambridge University Press.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, *58*(2), 203–210.

Fisher, R., Cavanagh, J., & Bowles, A. (2011). Assisting transition to university: Using assessment as a formative learning tool. *Assessment & Evaluation in Higher Education*, *36*(2), 225-237.

Folwer, F. J. (2009). *Survey research methods* (4th ed.). Beverly Hills, CA: Sage.

Freeman, D. (1993). Renaming experience, reconstructing practice. *Teaching and Teacher Educa*tion, *9*(5-6), 485-497.

Freeman, D. (2002). The hidden side of the work: Teacher knowledge and learning to teach. *Language Teaching, 35*, 1-13.

Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Glaser, B. G. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. Mill Valley, CA: Sociology Press.

Glaser, B. G. (1992). *Basics of grounded theory analysis: Emergence vs forcing*. Mill Valley, CA: Sociology Press.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago, IL: Aldine Publishing Co.

Goldstein, L. M. (2005). *Teacher written commentary in second language writing classrooms*. Ann Arbor, MI: The University of Michigan.

Goos, M., & Hughes, C. (2010). An investigation of the confidence levels of course/subject coordinators in undertaking aspects of their assessment responsibilities. *Assessment & Evaluation in Higher Education*, *35*(3), 315-324.

Greeno, J. G. (1998). The situativity of knowing, learning and research. *American Psychologist*, *53*, 5-26.

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology*, *39*, 281–291.

Gross, J. J. (2008). Emotion regulation. In M. Lewis, J. M. Haviland-Jones, and L.F. Barrett (Eds.), *Handbook of emotions* (3rd ed.) (pp. 497-512). New York, NY: Guilford.

Groundwater-Smith, S., & Mockler, N. (2007). Ethics in practitioner research: An issue of quality. *Research Papers in Education*, *22*(2), 199-211.

Haastrup, K. (1987). Using thinking aloud and retrospection to uncover learners' lexical inferencing procedures. In C. Færch & G. Kasper (Eds), *Introspection in second language research* (pp. 197-212). Clevedon, UK: Multilingual Matters.

Hamp-Lyons, L. (1991). Scoring procedures for ESL contexts. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts* (pp. 241 - 276). Norwood, NJ: Ablex.

Hamp-Lyons, L. (1996). The challenge of second-language writing assessment. In E. M. White, W. D. Lutz & S. Kamusikiri (Eds.), *Assessment of writing: Politics, policies, practices* (pp. 226-240). New York: Modern Language Association of America.

Hamp-Lyons, L. (1997). Washback, impact and validity: Ethical concerns. *Language Testing*, *14*(3), 295-303.

Hamp-Lyons, L. (2008). Writing assessment: Expanding outwards and coming together. *Assessing Writing*, *13*(1). pp. 1-3.

Hativa, N., Barak, R., & Simhi, E. (2001). Exemplary university teachers: Knowledge and beliefs regarding effective teaching dimensions and strategies. *The Journal of Higher Education*, *72*(6), 699-727.

Hattie, J., & Timperlay, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112.

Hawe, E. (2003). It's pretty difficult to fail: The reluctance of lecturers to award a failing grade. *Assessment and Evaluation in Higher Education*, *28*(4), 371-382.

Hood, M. (2009). Case study. In J. Heigham & R. A. Croker (Eds), *Qualitative research in applied linguistics: A practical introduction* (pp. 66-90). Basingstoke, UK: Palgrave Macmillan.

Hounsell, D., McCune, V., Hounsell, J., & Litjens, J. (2008). The quality of guidance and feedback to students. *Higher Education Research and Development*, *27*(1), 55-67.

Humphrey, N., Curran, A., Morris, E., Farrell, P., & Woods, K. (2007). Emotional intelligence and education: A critical review. *Educational Psychology*, *27*(2), 235-254.

Huot, B. A. (1993). The influence of holistic scoring procedures on reading and rating student essays. In M. A. Williamson & B. A. Huot (Eds.), *Validating holistic scoring for writing assessment* (pp. 207-236). Cresskill, NJ: Hampton Press.

Hutchins, E. (1995). How a cockpit remembers its speed. *Cognitive Science*, *19*, 265-288.

Imai, Y. (2010). Emotions in SLA: New insights from collaborative learning for an EFL classroom. *The Modern Language Journal*, *94*(2), 278–292.

Ivanic, R., Clark, R., & Rimmershaw, R. (2000). What am I supposed to make of this? The messages conveyed to students by tutors' written comments. In I. M. R. Lea & B. Stierer (Eds), *Student writing in higher education: New contexts* (pp. 47-65). Buckingham: SRHE & Open University.

Jacobs, A. (1974). The use of feedback in groups. In A. Jacobs & W. E. Spradlin (Eds.), *The Group as an Agent of Change* (pp. 408-448). New York: Behavioral Publications.

James, D. (2000). Making the graduate: Perspectives on student experience of assessment in higher education. In A. Filer (Ed.), *Assessment: Social practice and social product* (pp. 151-167). London, UK: RoutledgeFalmer.

Jawitz, J. (2009). Academic identities and communities of practice in a professional discipline. *Teaching in Higher Education*, *14*(3), 241-251.[not cited in text]

Jeffery, F., & Selting, B. (1999). Reading the invisible ink: Assessing the responses of non-composition faculty. *Assessing Writing*, *6*(2), 179-197.

Johnson, K. E. (1994). The emerging beliefs and instructional practices of preservice English as a second language teachers. *Teaching & Teacher Education*, *10*(4), 439-452.

Johnson. J. M. (2002). In-depth interviewing. In J. F. Gubrium & J. A. Holstein (Eds), *Handbook of interview research: context & method* (pp. 103-120). Thousand Oaks, CA: Sage.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed method research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14-26.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*, 130-144.

Joughin, G. (2009). Assessment, learning and judgement in higher education: A critical review. In G. Joughin (Ed.), *Assessment, learning and judgement in higher education* (pp. 13-27). Dordrecht: Springer.

Kagan, D. M. (1990). Ways of evaluating teacher cognition: Inferences concerning the Goldilocks principle. *Review of Educational Research*, *60*(3), 419- 469.

Kamberrelis, G., & Dimitriadis, G. (2005). Focus groups: Strategic articulations of pedagogy, politics, and inquiry. In N. Denzin & Y. Lincoln (Eds), *The Sage handbook of qualitative research* (3rd ed., pp. 887-907). London, UK: Sage.

Kane, R., Sandretto, S., & Heath, C. (2002). Telling half the story: A critical review of research on the teaching beliefs and practices of university academics. *Review of Educational Research*, *72*(2), 177-228.

Keh, C. (1990). Feedback in the writing process: A model and methods for implementation. *ELT Journal*, *44*(4), 294-304.

Kelle, U. (2005). "Emergence" vs. "Forcing" of empirical data? A crucial problem of "Grounded Theory" reconsidered. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *6*(2), Art. 27. doi:http://nbn-resolving.de/urn:nbn:de:0114-fqs0502275.

Kember, D., & Kwan, K. P. (2000). Lecturers' approaches to teaching and their relationship to conceptions of good teaching. *Instructional Science*, *28*(5), 469-490.

Knight, P., Tait, J., & Yorke, M. (2006). The professional learning of teachers in higher education. *Studies in Higher Education*, *31*(3), 319-339.

Kvale, S. (2007). Contradictions of assessment for learning in institutions of higher learning. In D. Boud and N. Falchikov (Eds), *Rethinking assessment*

in higher education: Learning for the longer term (pp. 57-71). New York, NY: Routledge.

Lantolf, J. P. (2006). Sociocultural theory and L2: State of the art. *Studies in Second Language Acquisition*, *28*, 67–109.

Lantolf, P., & Appel, G. (1994). *Vygotskian approaches to second language research*. Norwood, NJ: Ablex.

Lave, J. (1988). *Cognition and practice*. Cambridge, Uk: Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York, NY: Oxford Press.

Lea, M. R., & Street, B. V. (2000). Student writing and staff feedback in higher education: An academic literacies approach. In M. R. Lea & B. Stierer (Eds), *Student writing in higher education: New contexts* (pp. 31-46). Buckingham: SRHE & Open University Press.

Leary, M. R. (2000). Affect, cognition, and the social emotions. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 331-356). New York, NY: Cambridge University Press.

Lee, G., & Schallert, D. L. (2008). Meeting in the margins: Effects of the teacher-student relationship on revision processes of EFL college students taking a composition course. *Journal of Second Language Writing*, *17*, 165-182.

Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal*, *63*(1), 13-22.

Leki, I. (1990). Coaching from the margins: Issues in written response. In B. Koll (Ed.), *Second language writing: Research insights for the classroom* (pp. 57-68). Cambridge, UK: Cambridge University Press.

Leont'ev, A. N. (1978). *Activity, consciousness, and personality*. Englewood Cliffs, NJ: Prentice-Hall.

Li, J. R. (2009). Exploring teacher/tutor cognition with activity theory [Electronic version]. *TKKA, 4*.

Li, J. R., & Barnard, R. (2009). Differences of opinion: Methodological considerations regarding addressivity in individual interviews and focus groups. *New Zealand Studies in Applied Linguistics*, *15*(2), 15-29.

Li, J. R. & Farrell, T. (in press). Think aloud. In R. Barnard & A. Burns (Eds.), *Researching language teacher cognition and practice: International case studies*. Bristol, England: Multilingual Matters.

Lilly, J., Richter, U. M., & Rivera-Macias, B. (2010). Using feedback to promote learning: Student and tutor perspectives. *Practitioner Research in Higher Education*, *4*(1), 30-40.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. London, UK: Sage.

Lortie, D. (1975). *Schoolteacher. A sociological study*. Chicago, IL: University of Chicago Press.

Lynne, J. M. (2006). Focus groups. In G. M. Breakwell, S. Hammond & C. Fife-Schaw (Eds), *Research methods in psychology* (3rd ed., pp. 274-298). London, UK: Sage.

Mahn, H., & John-Steiner, V. (2002). The gift of confidence: A Vygotskian view of emotions. In G. Wells & G. Claxton (Eds), *Learning for life in the twenty-first century: Sociocultural perspectives on the future of education* (pp. 46-58). Oxford, UK: Blackwell.

Martin, A. K., Munby, H., & Hutchinson, N. L. (2000). Protests and praise from the field: Focus groups and predictive validity. In R. Upitis (Ed.), *Who will*

teach? A case study of teacher education reform (pp. 279-297). San Francisco, CA: Caddo Gap Press.

Mathison, S. (1988). Why triangulate? *Educational Researcher, 17*(2), 13-17.

McGee, S. J. (1999). *A qualitative study of student response to teacher-written comments.* Dissertation: Purdue University.

Mckay, S. L. (2009). Introspective techniques. In J. Heigham & R. A. Croker (Eds), *Qualitative research in applied linguistics: A practical introduction* (pp. 220-241). Basingstoke, UK: Palgrave Macmillan.

McNamara, T. F., & Roever, C. (2006). *Language testing: The social dimension*. Malden, MA: Blackwell.

Melzer, D. (2009). Writing assignments across the curriculum: A national study of college writing. *College Composition and Communication*, *61*, 240-261.

Merriam, S. B. (1998). *Qualitative research and case study applications in education.* San Francisco, CA: Jossey-Bass Publishers.

Merriam, S. B. (2002). *Qualitative research in practice: Examples for discussion and analysis*. San Francisco, CA: Jossey-Bass.

Meyer, L., Davidson, S., McKenzie, L., Rees, M., Anderson, H., Fletcher, R., & Johnston, P. (2010). An invesigaion of tertiary assessment policy and practice: Alignment and contradictions. *Higher Education Quarterly*, *64*(3), 331-350.

Moore, C., O' Neill, P., & Huot, B. (2009). Creating a culture of assessment in writing programs and beyond. *College Composition and Communication*, *61*(1), 107-132.

Moore, J. L., & Rocklin, T. R. (1998). The distribution of distributed cognition: multiple interpretations and uses. *Educational Psychology Review*, *10*(1), 97-113.[not cited in text]

Morgan, D. L. (2002). Focus group interviewing. In J. F. Gubrium & J. A. Holstein (Eds.), *Handbook of interview research: Context & method* (p. 141-159). Thousand Oaks, CA: Sage.

Morgan, D. L. (1997). *Focus group as qualitative research* (2nd ed.). London, UK: Sage.

Mutch, A. (2003). Exploring the practice of feedback to students. *Active Learning in Higher Education*, *4*(1), 24-38.

Nardi, B. A., (1996). Studying context: A comparison of activity theory, situated action models and distributed cognition. In B. Nardi (Ed.) *Context and Consciousness:Activity theory and human computer interaction* (pp. 69-103). Boston, MA: MIT Press.

Nicol, D. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. *Assessment & Evaluation in Higher Education*, *35*(5), 501-517.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218.

North, S. (2005). Disciplinary variation in the use of theme in undergraduate essays. *Applied Linguistics*, *26*(3), 431-452.

Nuan, D. (1992). *Collaborative language learning and teaching*. Cambridge, UK: Cambridge University Press.

Oatley, K. (1992). *Best laid schemes: The psychology of emotions*. New York, NY: Cambridge University Press.

Orrell, J. (2006). Feedback on learning achievement: Rhetoric and reality. *Teaching in Higher Education, 11*(4), 441-456.

Orsmond, P., & Merry, S. (2011). Feedback alignment: Effective and ineffective links between tutors' and students' understanding of coursework feedback. *Assessment & Evaluation in Higher Education*, *36*(2), 125-136.

Parker, A. & Tritter, J. (2006). Focus group method and methodology: current practice and recent debate. *International Journal of Research & Method in Education*, *29*(1), 23- 37.

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage.

Poehner, M. E. & Lantolf, J. P. (2010). Vygotsky's teaching-assessment dialectic and L2 education: The case for dynamic assessment. *Mind, Culture, and Activity*, *17*(4), 312-330.

Powell, R. P. (2002). The influence of prior experiences on pedagogical constructs on tradtional and non-traditional preservice teachers. *Teaching & Teacher Education*, *8*(3), 225-238.

Price, M., Carroll, J., O' Donovan, B., & Rust, C. (2011). If I was going there I wouldn't start from here: A critical commentary on current assessment practice. *Assessment & Evaluation in Higher Education*, *36*(4), 479-492.

Price, M., Handley, K., Millar, J., & O'Donovan. (2010). Feedback: All that effort, but what is the effect? *Assessment & Evaluation in Higher Education*, *35*(3), 277–289.[not cited in text]

Pula, J. J., & Huot, B. A. (1993). A model of background influences on holistic raters. In M. M. Williamson & B. A. Huot (Eds), *Validating holistic scoring for writing assessment: Theoretical and empirical foundations* (pp. 237-265). Cresskill, NJ: Hampton Press.

Quinlan, K. M. (1999). Commonalities and controversy in context: A study of academic historians' educational beliefs. *Teaching and Teacher Education*, *15*, 447-463.

Race, P. (1995). The art of assessing. *New Academic*, *4*(2), 3-6.

Read, B., Francis, B. & Robson, J. (2005). Gender, bias, assessment and feedback: Analyzing the written assessment of undergraduate history essays. *Assessment and Evaluation in Higher Education*, *30*(3), 241-260.

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, *35*(4), 435-448.

Richards, J. C. (1996). Teachers' maxims in language teaching. *TESOL Quarterly*, *30*(2), 281-296.

Rogers, Y. & Ellis, J. (1994). Distributed Cognition: An alternative framework for analysing and explaining collaborative working. *Journal of Information Technology*, 9(2), 119-128.

Roth, W. M. (2007). Emotion at work: A contribution to third-generation cultural-historical activity theory. *Mind, Culture, and Activity*, *14*(1-2), 40 - 63.

Russell, D. R. (1997). Writing and genre in higher education and workplaces. *Mind, Culture, and Activity, 4*, 224–237.

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education*, *5*(1), 77–84.

Sakyi, A. A. (2000). Validation of holistic scoring for ESL writing assessment: How raters evaluate compositions. In J. J. Kunnan (Ed.), *Fairness and validation in language assessment* (pp. 129-152). Cambridge, UK: Cambridge University Press.

Salomon, G. (1993). *Distributed cognitions: Psychological and educational considerations*. New York, NY: Cambridge University Press.

Samuelowicz, K., & Bain, J. (2002). Identifying academics' orientations to assessment practice. *Higher Education*, *43*(2), 173-201.

Scarino, A. (2005). Introspection and retrospection as windows on teacher knowledge, values, and ethical dispositions. In D. J. Tedick (Ed.), *Second language teacher education: International perspectives* (pp. 33-52). Mahwah, NJ: Lawrence Erlbaum Associates.

Schon, D. A. (1991). *The reflective turn: Case studies in and on educational practice*. New York, NY: Columbia University Press.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4-14.

Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review*, *74*, 29-39.

Sitko, B. (1993). Exploring feedback: Writers meet readers. In A. Penrose & B. Sitko (Eds), *Hearing ourselves think: Cognitive research in the college writing classroom* (pp. 170-187). New York, NY: OUP.

Smagorinsky, P. (1998). Thinking and speech and protocol analysis. *Mind, Culture, and Activity, 5*, 157–177.

Smith, C., & Bath, D. (2004). Evaluation of a university-wide strategy providing staff development for tutors: Effectiveness, relevance and local impact. *Mentoring and tutoring*, *12*(1), 107-122.

Salomon, G. (1993). *Distributed cognitions: Psychological and educational considerations*. New York, NY: Cambridge University Press.[not cited in text]

Stake, R. (1980). The case method inquiry in social inquiry. In H. Simons (Ed.), *Towards a science of the singular* (pp. 62-75). Columbus, OH: Merrill.

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.

Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed., pp. 443-466). Thousand Oaks, CA: Sage.

Stern, L. A., & Solomon, A. (2006). Effective faculty feedback: The road less traveled. *Assessing Writing*, *11*(1), 22-41.

Stewart, D. W., Shamdasani, P. N., & Rook, D. W. (2007). *Focus groups: Theory and practice.* Thousand Oaks, CA: Sage.

Storms, B., & Sheingold, K. (1999). *Teachers' and students' uses of and reactions to the DWS prototype*. Princeton, NJ: Educational Testing Service.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.

Sutherland, K. (2009). Nurturing undergraduate tutors' role in the university teaching community. *Mentoring & Tutoring: Partnership in Learning*, *17*(2), 147-164.

Sutton, R. E., & Wheatley, K. F. (2003). Teachers' emotions and teaching: A review of the literature and directions for future research. *Educational Psychology Review*, *15*(4), 327-358.

Swain, M. (2006). Verbal protocols. In M. Chalhoub-Deville, C. A. Chapelle & P. Duff (Eds), *Inference and generalizability in applied linguisitcs: Multiple perspectives* (pp. 97-114). Amsterdam, The Netherlands: John Benjamins.

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge, UK: Cambridge University Press.

Tang, J., & Harrison, C. (2010). Investigating university tutor perceptions of assessment feedback: Three types of tutor beliefs. *Assessment & Evaluation in Higher Education*, 1-22.

Taras, M. (2006). Do unto others or not: Equity in feedback for undergraduates. *Assessment & Evaluation in Higher Education*, *31*(3), 365–377.

Thompson, M. P. A. (2004). Some proposals for strengthening organizational activity theory. *Organization*, *11*(5), 579–602

Trigwell, K., & Prosser, M. (1996). Changing approaches to teaching: A relational perspective. *Studies in Higher Education*, *21*, 275-284.

Trope, Y., Ferguson, M., & Raghunathan, R. (2001). Mood as a resource in processing self-relevant information. In J. P. Forgas (Ed.), *Handbook of affect and social cognition* (pp. 256–274). London, England: Erlbaum.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, *46*(2), 327-369.

Truscott, J. (1999). The case for "The case against grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing*, *8*(2), 111-122.

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, *16*(4), 255-272.

Van Lier, L. (1996). *Interaction in the language curriculum: Awareness, autonomy and authenticity*. New York, NY: Longman.[not cited in text]

Värlander, S. (2008). The role of students' emotions in formal feedback situations. *Teaching in Higher Education*, *13*(2), 145-156.

Vaughan, C. (1991). Holistic assessment: What goes on in the rater's mind? In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts* (pp. 111-125). Norwood, NJ: Ablex.

Volante, L., & Fazio, X. (2007). Exploring teacher candidates'assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, *30*(3), 749-770.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: The MIT Press.

Vygotsky, L. S. (1978). *Mind in Society* (M. Cole, Trans.). Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1981). The instrumental method in psychology. In J. Wertsch (Ed.), *The concept of activity in Soviet psychology*. Armonk, NY: Sharpe.

Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: The MIT Press.

Vygotsky, L. S. (1987). Thinking and speech (N. Minick, Trans.). In R. W. Rieber and A. S. Carton (Eds.), *The Collected Works of L. S. Vygotsky: Vol. 1. Problems of general psychology* (pp. 39–285). New York: Plenum Press.

Weigle, S. C. (1999). Investigating rater/prompt interaction in writing assessment: Quantitative and qualitative approaches. *Assessing Writing*, *6*(2), 145-178.

Weigle, S. C. (2007). Teaching writing teachers about assessment. *Journal of Second Language Writing*, *16*(3), 194-209.

Wells, G. (1999). *Dialogic inquiry: Towards a sociocultural practice and theory of education*. NY: Cambridge University Press.

Wilson, T. D. (1994). The proper protocol: Validity and completeness of verbal reports. *Psychological Science, 5*, 249-252.

Winter, J. K., Neal, J. C., & Waner, K. K. (1996). Student and instructor use of comments on Business Communication papers. *Business Communication Quarterly*, *59*(4), 56-68.

Woods, D. (1996). *Teacher cognition in language teaching: Beliefs, decision-making, and classroom practice*. Melbourne, Australia: Cambridge University Press.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Child Psychiatry, 17*, 89-100. [not cited in text]

Yin. R. K. (1994). *Case study research* (2nd ed.). Beverly Hills, CA: Sage.

Yin. R. K. (2003). *Case study research: design and methods* (3rd ed.). Thousand Oaks, CA: Sage.

Yin. R. K. (2006). Case study methods. In J. L. Green, G.Camilli & P. B. Elmore (Eds.), *Handbook of complementary methods in education research* (pp. 111-122). Mahwah, N.J.: Lawrence Erlbaum.

Zhu, J., & Thagard, P. (2002). Emotion and action. *Philosophical Psychology*, *15*(19-36).

APPENDICES

LIST OF APPENDICES

Appendix A: Ethical approval

Appendix B: Cover page of the e-questionnaire

Appendix C: E-Questionnaire for the tutors

Appendix D: Letter of informed consent to tutors

Appendix E: Focus points of the semi-structured interview

Appendix F: E-mail to tutors about the think-aloud and stimulated recall sessions

Appendix G: Think-aloud and stimulated recall instructions

Appendix H: Field notes in think-aloud sessions example

Appendix I: Topics for focus group discussion

Appendix J: Research journal example

Appendix A: Ethical approval

FASS HUMAN RESEARCH
ETHICS COMMITTEE   Phone +64 7 856 2889
The University of Waikato   www.waikato.ac.nz
Private Bag 3105
Hamilton, New Zealand

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

Jinrui Li
Dr Roger Barnard
Dr Maria Galikowski
Dr Rosemary De Luca

10 September 2008

Dear Jinrui Li

Application for Ethical Approval: **Convergence and Divergence of University Teachers' Beliefs and
Practices…**

Thank you for submitting the revisions to your Application for Ethical Approval in response to my letter of 8
September. Thanks also for meeting with me in the morning of 10 September to discuss the issue of
anonymity. Your revisions were received on the afternoon of that day, today.

I have considered your revisions to the section of the Application on anonymity and to the Information
Letters and Consent Forms. Thank you for responding so clearly and constructively to the suggestions I
had made.

I would also like to acknowledge the fine support and thoughtful advice you are receiving from Dr Barnard.

I am pleased to provide formal ethical approval for your research project.

I wish you a rewarding and successful outcome.

With best wishes,

John Paterson
Chair
FASS Human Research Ethics Committee

Appendix B: Cover page of the e-questionnaire

Doctoral Research Project

Teachers' Beliefs and Practices of Giving Feedback on Students' Written Assignments

Hello,
I hope you will agree to take part in the survey, which will only take about 10 to
15 minutes of your time.

The survey has been approved by FASS Human Research Ethics Committee.
If you consent to take part, please fill in the radio buttons after each item on the
questionnaire.

If you wish to make any changes, click 'reset' before submitting.

When you have completed the questionnaire to your satisfaction, please click 'submit'.
Your answers will remain anonymous, and your privacy and confidentiality will be
respected at all times.

I would very much appreciate it if you would submit the completed questionnaire before
Oct. 22nd.

If you have any questions on the survey, or on the research project of which this is part,
please email either myself (jl287@students.waikato.ac.nz) or my Chief Supervisor, Dr
Roger Barnard, (rbarnard@waikato.ac.nz).

Thank you very much for your help and attention!

Jinrui Li
PhD Student
Department of General and Applied Linguistics

(The web-link: www.waikato.ac.nz/wfass/survey/assignment-feedback/)

Appendix C: E-Questionnaire for the tutors

1．I work as a tutor in the Department of:	

	Less than one year	1-3 years	4-5 years	6 or more years
2. I have been tutoring in FASS for:	●	○	○	○

3a. I worked for approximately _____ hours as a tutor in Semester A, 2008.	

3b. I worked for approximately _____ hours as a tutor in Semester B, 2008.	

4．I have approximately _____ undergraduate students in my tutorial class. (Please specify the number. If you have more than one tutorial class, please respond to this question with one of your classes in mind.)	

4b. I have approximately _____undergraduate **international** students in my tutorial class. (Please specify the number. If you have more than one tutorial class, please respond to this question with one of your classes in mind.)	

Questions 5-9 consist of a series of statements. Which of the statements applies to you.

5. As a tutor, I:		
	Yes	No
Provide specific guidelines for completing written assignments	○	●
Discuss lecturers' guidelines in class before students submit assignments	●	○
Discuss lecturers' guidelines with individual students prior to submission of assignments	●	○
Review drafts of students' assignments and provide feedback	●	○
Mark written assignments and provide written feedback	●	○

	Yes	No
Mark written assignments without providing written feedback	⦿	○
Discuss feedback with each individual student	⦿	○
Discuss feedback with individual students where there is a particular issue with his/her assignment.	⦿	○
Provide feedback to students during group tutorials	⦿	○

6. The guidelines I receive from my lecturer include:

	Yes	No
Written instructions	○	○
Checklists	⦿	○
Listed criteria	⦿	○
One-to –one verbal instruction	⦿	○
Verbal instruction given to tutors in groups	⦿	○
Sample feedback	⦿	○
Model answers	⦿	○
Other (please specify)		

7. The written assignments I tutor cover:

	Yes	No
Short written work (one or two paragraphs)	○	○
Summaries of readings materials	⦿	○
Reviews of published work	⦿	○
Literature reviews	⦿	○
Essays	⦿	○
Reports	⦿	○

8. In my feedback, I focus on:		
Grammatical errors, such as spelling, choice of words, sentence structure, etc.	Yes ○	No ○
Subject content (content related to a discipline area)	●	○
Organization or structure at paragraph level	●	○
Organization or structure at text level	●	○
Referencing	●	○
Other (please specify)		

9. With regard to giving written feedback:		
I spend more than half an hour on each student assignment.	Yes ○	No ○
I red-mark errors in students' assignments to bring these errors to students' attention.	●	○
I underline/circle errors without giving corrections.	●	○
I underline/circle errors and give detailed corrections.	●	○
I selectively correct the errors students make in their assignments.	●	○
The only written feedback I provide is Comments at the end of students' assignments.	●	○
I do not give written feedback, and only provide a grade/mark on assignments.	●	○

Question 10 consists of a series of statements. Please indicate your agreement. **A non-response will indicate 'No Opinion'**

	Strongly agree	Agree	Disagree	Strongly disagree
Giving written feedback is time consuming and ineffective	●	○	○	○
Giving written feedback is important and effective	○	○	○	○

208

Native English speaker students have fewer problems in their writing than non native speaker students	Strongly agree ○	Agree ○	Disagree ○	Strongly disagree ○
Most students will not read feedback carefully	Strongly agree ○	Agree ○	Disagree ○	Strongly disagree ○
Please feel free to add any further relevant comments regarding feedback on students' assignments.				

Reset Submit

Appendix D: Letter of informed consent to tutors for follow-up sessions

(A case study of feedback given on the written assignments of undergraduate students in a New Zealand university)

Dear_____
Thank you for agreeing to participate in follow-up sessions. As you know, I am interested in identifying and exploring teachers' beliefs and practices about feedback given on the written assignments of undergraduate students.
Following the survey and our recent contact, I shall now like you to take part in a series of follow-up sessions during semester A of 2009.

First, at the beginning of the semester, I shall invite you to take part in a semi-structured interview, during which I would ask you to respond (as fully as you feel able) to four or five questions relating to your attitudes and practices about feedback given to students' assignments.

Second, I should like to be present when you are drafting written feedback on your students' assignments. I shall invite you to "think aloud", which means you will talk through your actual activity while giving feedback, with or without some limited verbal cues provided by me. If feedback is given in the form of individual or group conference, I shall present as a non-participant observer to record your oral feedback.

Third, I shall invite you to meet with me after you will have given feedback on your students' assignments, during which I shall ask you some questions contingent on the responses you will have made during the previous sessions.
Each interview or session will last between thirty and sixty minutes and each interview or session will be audio-taped. I shall send you a summary of each interview or session after that.

I should like to assure you that the research will adhere strictly to the University of Waikato Human Research Ethics Regulations (2008). Your right to anonymity and privacy will be respected during and after the research process. No real names will be used in the research report, and all the data gathered will be kept confidential. The anonymized interview data will only be seen by myself and my supervisors, and care will be taken to ensure that no individual can be identified from any resulting report or publication. Any information gathered will only be used for the academic purposes of this research thesis or any resulting journal or conference presentations, unless your permission is obtained for other uses. All information will be coded and the information gathered will have no negative impact on your current work at the University of Waikato. All data gathered in the research will be stored in a locked cupboard in my office in the Department of General and Applied Linguistics, to which only the chief supervisor and I have access. It will be kept safe there for the minimum of five years for the purpose of academic examinations and reviews. After that, all the research data will be destroyed. Please note that you may withdraw your participation from the project at any time, with no need to give any reason for so doing.

Your participation will be greatly appreciated. If you are willing to take part in these interviews, please complete the consent form below, and return it to me by email, to my office (J3.14), or I can collect it from your office if advised. It will be useful for you to keep a copy of this letter and the form for your personal records.

This project has been approved by the Human Ethics Research Committee of the University of Waikato, and any questions regarding the ethical conduct of this project may be addressed to the Secretary of the Committee (fass-ethics@waikato.ac.nz). Of course, if you have further enquires about the project, please contact me directly by cellphone: 0211074386, telephone: 07 838 4466, extension 6777, e-mail: jl287@students.waikato.ac.nz, or contact my chief supervisor Dr. Roger Barnard by telephone: 07 8384466, extension 6691, Email: rbarnard@waikato.ac.nz.

With many thanks,
Jinrui Li

Appendix E: Focus points of the semi-structured interview

1. What do you believe is good writing for your students?
2. Do you discuss criteria with your students before they submit the assignment?
3. Do you give students written feedback on their assignments?
4. What is your purpose of giving feedback?
5. What do you think is a good practice of giving feedback?
6. Could you tell me the focus of your feedback? Why?
7. Do you use any strategies or techniques while giving feedback?
8. Do you give feedback on drafts?
9. How much time do you spend on each assignment?
10. Do you think native speaker students and non-native speaker students need the same kind of feedback?
11. Do you think your feedback is effective? Why?
12. Do you receive students' feedback on your feedback?
13. Could you tell me the sources of this belief?
14. Do you have any other comments on feedback?

Appendix F: E-mail to tutors about the think-aloud and stimulated recall sessions

Dear [NAME],

Thank you very much for allowing me to be present when you give feedback on one or two of your students' assignments. What I wish you to do is to talk aloud everything that goes through your mind when you are giving feedback on the written assignments.

While you are thinking aloud, I will remain silent, and perhaps make some notes. I may also prompt you to say a little more about what you are thinking.
I wish to do a stimulated recall with you after the think aloud session. Please advise me which time span you prefer between think aloud and stimulated recall: 1) immediately after the think aloud 2) half or an hour later 3) several hours later 4) a day later.

During the stimulated recall, we are going to listen to the tape of the think aloud session. I am interested in what you were thinking at the time you were giving the feedback. What I'd like you to do is to tell me what you were thinking, what was in your mind at that time while you were talking and giving feedback.

Appendix G: Think-aloud and stimulated recall instructions

Oral instructions for think-aloud sessions

Thank you very much for your participation in the think-aloud session. What I wish you to do is to talk aloud everything that goes through your mind when you are giving feedback on the written assignments. While you are thinking aloud, I will remain silent and perhaps make some notes. I may also prompt you to say a little more about what you are thinking.

Prompt: Could you please tell me what you are thinking about?

Oral instructions for stimulated recall sessions

I am interested in what you were thinking at the time you were giving the feedback. I could hear what you were saying but I did not know what you were thinking. So what I'd like you to do is to tell me what you were thinking. What was in your mind at that time while you were talking?

Questions:

What: While giving feedback, you did…, what were you thinking about?
Why: While giving feedback, you said… Why did you say so?

Appendix H: Field notes in think-aloud sessions example

1. pencil →.
2. report.
3. circl the 1st ⊙typical⊙ example. ⌐ sel. of   ⌐ L2

   Difficulty in
   marking 2nd layut
   _____ lennere ?

4. laugh.. explain why you said this _____ L3.

   ̸.

   ̸ . laugh again.

   .. 1Definition. guesture. (hands)

5. .. been) comment . phrasing ⟨academ way⟩ ⟩ ⌐ L4

6. inclus. ~~~~ erase sth. ◁

⟨Sally negative⟩ erase again
   .. _____ cross a part of senteme : _____ L6
   put the correction above.
8. crossout ~~paint~~ word and then read & ~~erase~~ the crossing add "four". ⌐

8. cannot grade 2nd language Ss in this respect . ⟨C8⟩

   {   ) I don't think that is necessary . so . if you don't argue :
   ... comment on the paraa. on the left margin _____ C7.

   (It seems the T present his cognitive process of reading on the assignment
   a dialogue without thinking to much of ~~pedagogy~~ – it is a equal posirion with Ss.)

8 lines in the 1st par.
   Cross most lines . change (   ). _____ L8 ⟨
   Erase again . ~~coffee~~
   (write down a comment at the end of the 1st paxi. _____ C2
   ~~connect~~ content right hand of the margin   ⟨C2⟩
   ?   -x-x-^.. ↑
   _____ L9   . ⌐ L10
Para.2. ~~1 mainly focus~~ 2 ̸ My mainly focus
   ~~My mains focus.~~

215

Appendix I: Topics for focus group discussions

Thank you very much for agreeing to participate in the focus group. The following are among the specific concerns commonly expressed:

Marking: Tutors tended to avoid giving extremely high or extremely low (especially 'fail') grades because they were aware, in the latter case, of the social and academic impact.

Empathy: As students themselves, tutors may be excessively empathetic to the students whose assignments they were marking.

Formal correction: Most tutors tended not to pay specific attention to either surface errors (syntax, spelling, etc) or to organisational/structural features, but to focus on the discipline-related content. Tutors felt that they lacked competence to correct these matters and, in some cases, referred the students to TLDU (or equivalent).

Feedback: While most tutors were encouraged to provide both positive and negative feedback, the amount of comment on each of these varied; sometimes four or five lines of summary comment were considered appropriate; in other cases, simply writing words like 'good' were deemed sufficient.

Referencing: While the criteria always mentioned these, there was a lack of clear understanding among the tutors about how closely the students should follow a particular format.

Plagiarism: There was a constant need to refer plagiarism issues to the senior colleague/lecturer, as strict adherence to the guidelines would suggest low or fail grades.

Timing: tutors were hourly-paid, and the above matters added to the time needed to 'mark' the assignments, but not to their pay–packet.

Inexperience: Tutors (especially new tutors) were aware of their lack of experience, and relied very heavily on the guidance of their lecturers or senior colleagues: sometimes this advice conflicted with their own judgment.

I would be very grateful if you could comment on (agree or disagree, and discuss change, add on, explain etc.) any of the above points which interest you. Your comments will be very valuable as you help me to have a better understanding of the issues. Thank you again for your help.

Appendix J: Research journal example

Reflection on the pilot think-aloud session

The day after interview, I had a think-aloud session with the participant. He asked me to come around 5pm but when I arrived, he had already started marking. He asked me to come in, sit down and read the assignments he had already marked, during which he didn't stop marking. Then he stopped and before he started to mark a new one, he showed me the criteria he used while marking. Then he showed me how he marked by sitting together with me and talked about his marking process. It seemed he wouldn't do a monologue by himself. He was more willing to tell me what he was doing. While I observed him marking the second assignment, I asked whether he could forget my existence and just think aloud. He remained silent for a while, tried to murmur for himself and then explained his operations to me again.

I realize that monologue or strict think-aloud is too experimental to be successfully used by my teacher participants. Tim preferred a dialogic way, even though I just showed my attention by nodding and back-channelling. It seemed talking to me was more natural for him than doing a monologue to himself. Compared with the previous think-aloud session while I sat in a corner of a room, observing my participant doing monologue, this way seems better for three reasons. Firstly, while sitting together with the participant, I can see clearly what he is writing and how he writes his feedback. Therefore, I can note down rich data by close observation. Secondly, my close observation its self is a prompt for participant. It motivates participant to talk about his operations because there is an audience in the room, which can be regarded as a natural way of thinking aloud, the participant feels more willing to provide data than in this way than in an experimental way. Third, the researcher can copy the feedback at the same time without asking for photocopies of the feedback. It is especially helpful when it is impossible to get copies of the marked assignments. Can think aloud be done in a natural way together with a close observation of the researcher? Yes. Can I make it natural by showing my attention of the talk while I am mainly a listener? I'll try.