

FEEDBACK IN TECHNOLOGY BASED INSTRUCTION: LEARNER PREFERENCES

DAVID JOHN LEFEVRE

The Business School, Imperial College London.

Thesis submitted for the degree of Doctor of Philosophy

July 2013

DECLARATION OF ORIGINALITY

I declare that the following work is my own unless indicated otherwise.

David Lefevre

COPYRIGHT DECLARATION

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

ACKNOWLEDGEMENTS

As anyone who has worked on their own doctoral research will understand, the journey to the completion of this thesis was long and sometimes arduous. I would like to express my sincere gratitude to my doctoral supervisor, Dr Benita Cox, for her tremendous patience and support throughout.

I would also like to thank my family for all their support. Particularly my wife, Akiko, who sacrificed so many weekends, evenings and holidays together to enable me to complete this work.

ABSTRACT

This research investigates learners' requirements for the format of feedback when using technology based instruction (TBI). A novel approach is taken by focusing on learners' preferences for such feedback. The primary method of data collection is to provide subjects with a range of options for TBI feedback and observe those options they choose to view.

Feedback has a powerful effect on learning. It is considered one of the ten most significant influences on learner achievement (Hattie and Timperley 2007: 83). Prior research into TBI feedback has focused on the relationship between the format of such feedback and scores achieved in subsequent performance tests. Results have been inconclusive and the format of TBI feedback that should be provided to learners is unclear. This has led to calls for an increased focus on the learners' perspective of TBI feedback (Mory 2003: 770; Handley, Price, et al. 2011: 548).

This research furthers understanding of the learners' perspective by identifying the format of TBI feedback learners choose to view and factors that affect this choice, hitherto an area not subject to empirical research. The specific focus is learners' preferences for TBI feedback provided following responses to multiple choice questions. A mixed-method approach is taken comprising technology based observations and semi-structured interviews. In the former, a software tool developed as part of this research delivers items of TBI to students and records their preferences for feedback. Between 2008 and 2010 data were collected on 24 cohorts of students studying one or more of three items of TBI. In total, students responded to 61,929 questions.

Results are interpreted through the constructivist perspective and the results are organised into eight themes which are then synthesised with reference to theoretical frameworks to produce a model of learner interaction with TBI feedback. In addition to furthering our understanding of how learners engage with TBI feedback, the results have implications for existing theoretical models and for research aiming to establish the effectiveness of different formats of TBI feedback. For the latter, this research also reveals avenues of future research that potentially address the inconsistent pattern of results found in prior research.

No claims are made as to the effectiveness of different feedback formats however a number of recommendations to educators regarding the content, timing and delivery of TBI feedback are suggested.

CONTENTS

Declaration of Originality, Copyright Declaration	3
Acknowledgements.....	5
Abstract.....	7
Contents.....	9
List of tables	10
List of figures	12
1. Introduction.....	15
1.1 Background.....	15
1.2 Defining TBI feedback.....	16
1.3 Importance of topic	17
1.4 Research questions.....	24
1.5 Summary of research method	24
1.6 Contribution to knowledge	25
2. Literature review	27
2.1 Procedure.....	27
2.2 Presentation of findings	29
2.3 The Research gap	54
2.4 Forming hypotheses	57
2.5 Literature review: Summary.....	63
3. Research methodology	65
3.1 Theoretical framework	65
3.2 Research design.....	71
3.3 Research methods	83
3.4 Key assumptions	100
3.5 The three items of Technology Based Instruction (TBI)	102
3.6 The development of the software tool.....	112
4. Research integrity.....	121
5. The four studies: Detailed design, results and analysis.....	125
5.1 Overview of the four studies.....	125
5.2 Study one: Establishing base data (2008).....	126
5.3 Study two: Qualitative interviews (2008)	148
5.4 Study three: Subsequent assessment and question focus (2009)	180
5.5 Study four: Factor ‘#I4 Level of expertise’ (2010).....	208
6. Summary of results.....	233
7. Discussion.....	249
7.1 The eight themes	249
7.2 A model of learner interaction with TBI feedback.....	260
8. Conclusions	265
8.1 Implications of the results	265
8.2 Critique of the design and conduct of the research	268
8.3 Limitations	271
8.4 Suggestions for future research.....	272
9. Appendices	275
Appendix one: Glossary.....	275
Appendix two: Recommendations for practitioners	278
Appendix three: Topic guide for semi structured interviews.....	281
Appendix four: Comparison of prior research on feedback taxonomy.....	283
Bibliography	285

LIST OF TABLES

Table 1: Key predictions from Kulhavy and Stock (1989).....	35
Table 2: Questions to be addressed by effective feedback (Hattie and Timperley 2007: 86).....	41
Table 3: Elaborative feedback types	41
Table 4: A taxonomy for the timing of TBI feedback	41
Table 5: A taxonomy of TBI feedback	46
Table 6: Descriptive statistics relating to the ' α ' values for each TBI feedback type.....	87
Table 7: Representation of attributes in the taxonomy of TBI feedback.....	111
Table 8: Factors that potentially affect whether a learner elects to view a feedback type	127
Table 9: Summary data relating to study one.....	130
Table 10: Summary statistics relating to α figures for each feedback type.....	130
Table 11: Relationship between IB and EX type feedback.....	132
Table 12: Relationship between KR and KCR type feedback.....	132
Table 13: Interplay between KR & KCR type feedback (Questions answered incorrectly	133
Table 14: Subjects who did not view the elaborative feedback types.....	135
Table 15: Data relating to delayed and immediate feedback.....	135
Table 16: VIF Factors	137
Table 17: Results of multilevel logistic regression analyses	138
Table 18: α values for each feedback type – Incorrect responses only.....	139
Table 19: α values for each feedback type – Correct responses only.....	139
Table 20: Number of questions revisited by subjects.....	140
Table 21: The nine potential factors identified in the literature review	154
Table 22: Summary of Handley's (2003: 212) styles of engagement	155
Table 23: The 13 possible factors for consideration for investigation in study three	182
Table 24: Question focus and subject matter of question in the three items of TBI.....	182
Table 25: Summary data relating to the cohorts in study three.....	185
Table 26: Summary data relating to TBI in study three	185
Table 27: α values for each feedback type.....	185
Table 28: Ethics and Strategy exercise - α values for each feedback type.....	188
Table 29: Pre-Study Mathematics course - α values for each feedback type.....	188
Table 30: Accounting Primer course - α values for each feedback type	188
Table 31: The interplay between IB and EX type feedback	189
Table 32: Relationship between KR and KCR type feedback.....	189
Table 33: α values for IB, EX, EI1 and EI2 type feedback.....	189
Table 34: Number of subjects who never viewed a feedback type	192
Table 35: Ethics and Strategy exercise only	192
Table 36: Accounting Primer course only	193
Table 37: Initial variance inflation factors	195
Table 38: Variance inflation factors following removal of 'programme' as a factor	195
Table 39: Results of univariate and multivariate multilevel logistic regression analyses	196
Table 40: Summary of results for factors significant to a p level of less than 0.05.....	198
Table 41: Subjects selected for analysis by cohort	199
Table 42: Subjects revisiting questions.....	199
Table 43: Mean α values for each feedback type by item of TBI	202
Table 44: Mean α values expressed as number of questions	202
Table 45: Mean α values for each feedback type – All items of TBI.....	202
Table 46: Mean α values for each feedback type – For each item of TBI.....	202
Table 47: The 13 possible factors for consideration for investigation in study four.....	209
Table 48: Summary data relating to the cohorts in 2010-2011	210
Table 49: The student cohorts forming the research subjects in study four.....	211

Table 50: Summary data relating to the two items of TBI used in study four	212
Table 51: α values for each feedback type.....	212
Table 52: α values for each feedback type – Accounting Primer course only.....	212
Table 53: α values for each feedback type – Pre-study Mathematics course only.....	213
Table 54: The interplay between IB and EX type feedback	213
Table 55: The interplay between KR and KCR type feedback	214
Table 56: Mean α values for IB, EX, EI1 and EI2 type feedback	214
Table 57: Accounting Primer Course only	217
Table 58: Pre-study Mathematics only	217
Table 59: Cohorts for whom timing data were gathered in the academic year 2010 to 2011.	218
Table 60: Feedback timing.....	219
Table 61: Variance inflation factors (VIFs)	220
Table 62: Results of univariate and multivariate multilevel logistic regression analyses	221
Table 63: Summary of results for factors significant to a p level of less than 0.05.....	224
Table 64: Cohorts selected for analysis for hypothesis 5.....	224
Table 65: Subjects revisiting questions.....	224
Table 66: Mean α values by item of TBI.....	226
Table 67: Mean α values –Cohorts studying Accounting Primer course in 2009 & 2010.....	227
Table 68: Mean α values –Cohorts studying Pre-study Math course in 2009 & 2010	228
Table 69: Odds ratios for the significant effects of factor ‘#E5 question focus’.....	229
Table 70: α values for each feedback type for all subjects across this series of studies.....	233
Table 71: α values for all subjects across studies one to four by cohort and item of TBI.....	233
Table 72: Use of KCR type feedback - Incorrect responses only.....	237
Table 73: Mean α values for feedback types IB, EX, EI1 and EI2 for all cohorts	236
Table 74: Preference for the timing of feedback.....	237
Table 75: Factors that potentially affect whether learners choose to view a feedback type	238
Table 76: The five factors investigated via logistic regression	238
Table 77: The ten factors investigated in the semi-structured interviews.....	238
Table 78: Significant effects for Factor #I5 Performance	242
Table 79: Significant effects for factor ‘#E1 assessment’.....	242
Table 80: Significant effects for #E5 Question focus.....	245
Table 81: Summary data relating to subjects revisiting questions	247

LIST OF FIGURES

Figure 1: The location of topic within educational research.....	17
Figure 2: Index of Student Numbers and Public Funding for Higher Education.....	20
Figure 3: The three-cycle feedback paradigm (Kulhavy and Stock 1989: 291)	34
Figure 4: A five-stage model - Bangert-Drowns, Kulik et al.....	36
Figure 5: Handley's conceptual framework (Handley 2003: 206)	39
Figure 6: A model of Feedback to Enhance Learning (Hattie and Timperley 2007: 87)	42
Figure 7: The four levels of complexity (Kulhavy, White et al. 1985: 287)	51
Figure 8: Kluger and DeNisi (1996) - Three levels of control.....	54
Figure 9: The timeline of this research.....	84
Figure 10: Frequency distribution of α values for the IB feedback	88
Figure 11: Log (left) and Square root (right) transformations	88
Figure 12: Inverse (left) and Square transformations	88
Figure 13: Data relating to incorrect responses	89
Figure 14: A screenshot showing the structure of the Accounting Primer course.....	104
Figure 15: A screenshot showing the structure of the Pre-study Mathematics course.....	105
Figure 16: Example of EX type feedback from the Ethics and Strategy exercise.	106
Figure 17: Example of (EX) type feedback from the Accounting Primer course.	106
Figure 18: Example of IB type feedback from the Ethics and Strategy Exercise.....	107
Figure 19: Example of IB type feedback from the Accounting Primer course.....	107
Figure 20: Example of IB type feedback from the Pre-study Mathematics course.....	108
Figure 21: Example of EI1 type feedback from the Accounting Primer course	108
Figure 22: Example of EI1 type feedback from the Pre-study Mathematics	109
Figure 23: Example of EI2 type feedback from the Ethics and Strategy exercise	109
Figure 24: Example of EI2 type feedback from the Accounting Primer course	110
Figure 25: Example of EI2 type feedback from the Pre-study Mathematics course	110
Figure 26: The registration screen.....	112
Figure 27: The question screen.....	112
Figure 28: The feedback screen.....	113
Figure 29: The feedback screen showing an example of IB type feedback	113
Figure 30: A screenshot of the reporting interface	113
Figure 31: A screenshot of the reporting interface	113
Figure 32: Different arrangements for the six feedback types.....	113
Figure 33: Inherent orders in the different arrangements	118
Figure 34: A screenshot of the interface presented to subjects upon answering a question	118
Figure 35: The timeline of this research.....	125
Figure 36: Distribution of α values for 'No feedback' and KR type feedback.....	125
Figure 37: Distribution of α values for KCR type feedback and subjects' scores.....	134
Figure 38: Distribution of α values for EX/IB/EI1/EI2 type feedback.....	134
Figure 39: An example of a subject pathway	134
Figure 40: Total clicks across all feedback types per question.....	134
Figure 41: Feedback clicks for KR type feedback.....	141
Figure 42: Feedback clicks for KCR type feedback.....	141
Figure 43: Feedback clicks for EX type feedback.....	141
Figure 44: Feedback clicks for IB type feedback	142
Figure 45: The distribution of α values for KCR, IB, EX, EI1 and EI2 type feedback	142
Figure 46: Histogram showing the distribution of performance scores (%) across subjects.....	192
Figure 47: The distribution of α values for KCR type feedback - Pre-study Mathematics	192
Figure 48: Feedback clicks for KR type feedback across the 'Ethics & Strategy' test	200
Figure 49: Feedback clicks for KR & KCR type feedback	200
Figure 50: Screenshot from the Accounting Primer course.....	204

Figure 51: Distribution of mean α values for KCR, IB, EX, EI1 and EI2 feedback types.....	215
Figure 52: The frequency distribution of performance scores (%) across all subjects.....	216
Figure 53: Dist. of α values for IB/EX feedback types - Pre-study Mathematics	216
Figure 54: Dist. of α values for IB/EX feedback types - Accounting Primer course	216
Figure 55: Performance scores and mean α values for IB feedback type in study 4	216
Figure 56: α values for IB type feedback for subjects in study four.....	216
Figure 57: An example of a subject' pathway	216
Figure 58: A model of the learner engagement with the TBI feedback types	216

1. INTRODUCTION

1.1 BACKGROUND

This research investigated learners' preferences for the format of feedback provided following responses to multiple choice questions within technology based instruction. Such feedback is termed 'TBI feedback' and for the purposes of this research, TBI feedback is defined as

Any communication or procedure given to inform a learner as to the accuracy of a response in the course of technology based instruction.

Educational feedback is one of the 5 to 10 most significant influences on student achievement (Hattie and Timperley 2007: 83) and a large body of research has focused on the role educational feedback plays in learning. Since the early 20th century, educational feedback has been a central component of educational theory (For example, Skinner 1938; Kolb and Fry 1975; Laurillard 2002). Instructional feedback, that is educational feedback provided within a sequence of instructional activities, comprises one of Robert Gagne's nine events of instruction (Gagne 1985), a widely adopted and frequently cited theory of instructional design (Driscoll 2000: 371).

An increase in the use of technology based instruction in the 1970s and 1980s prompted research into the nature of instructional feedback provided by computers. The majority of studies since this time have focused on the relationship between the format of TBI feedback and scores in subsequent performance tests (For example, Gilman 1969; Hanna 1976; Phye 1979; Andre and Thieman 1988; Kulhavy and Stock 1989; Pridemore and Klein 1995; Moreno 2004). Despite significant effort in this direction, periodic reviews of the literature express frustration at the results (Kulhavy and Stock 1989; Mory 2003; Shute 2008). The following statement by Shute is an example,

“Within this large body of feedback research, there are many conflicting findings and no consistent pattern of results” (Shute 2008: 153).

Novel approaches to TBI feedback research are required. Despite constructivism being the dominant paradigm in contemporary educational research, Mory's (2003: 770) observation that there had been little research into TBI feedback from a constructivist perspective is still accurate. Constructivist research emphasizes the role of the learner in the learning process and the strategies learners adopt when constructing knowledge and acquiring skills. This call for a focus on the learner is echoed by Handley, Price, et al. (2011: 548) who argue that the absence of consideration of learners' engagement with feedback explains the inconsistency found in prior research.

This research contributes towards addressing this gap by furthering our understanding of how learners engage with TBI feedback.

1.2 DEFINING TBI FEEDBACK

In this research ‘educational feedback’ refers to information provided by an agent to a student regarding their learning performance or level of understanding. The agent providing information can be any within the educational environment. For example the self, a teacher, a peer, a book or a computer. The nature of the educational feedback provided can range from a teacher assigning a grade to an assignment, to a fellow student sharing a model answer, to a parent arranging a personal tutor (Hattie and Timperley 2007: 81).

One type of educational feedback is ‘instructional feedback’ which is defined here as feedback provided in the course of instruction. Instruction being defined as,

“Any deliberate arrangement of events to facilitate a learner’s acquisition of some goal.” (Driscoll 2000: 25).

Within the context of instruction, feedback is generally provided following a learner’s response to an instructional event. A common definition is that adopted by Edna Mory in her 2003 literature review on the subject,

“(Instructional feedback is) any communication or procedure given to inform a learner of the accuracy of a response.” (Mory 2003 citing Carter, 1984; Cohen, 1985; Kulhavy, 1977; Sales, 1993: 745)

The focus of this research is one form of instructional feedback, namely that provided in the course of technology based instruction. This is termed ‘technology based instructional feedback’ or ‘TBI feedback’ in this research. That TBI feedback is provided by a computer and not a human teacher places potential restrictions on the form it can take. First, the content of the feedback is determined in advance of a learner’s response, as is the range of learner responses that will trigger the delivery of this feedback. Second, the nature of the feedback to be delivered is limited to the communicative qualities of a computer’s output devices, for example the monitor or audio speakers.

Definitions of TBI feedback have adopted a similar form in both structure and meaning to that of Mory (2003) above. For example, Wager and Wager (1985: 2) adopt an identical definition with the caveat that the feedback must be communicated via a computer display. Sales (1993) expands on this definition by adding a statement relating to the role played by TBI feedback,

“Information presented to the learner after any input. Its purpose is to shape the perceptions of the learner” (Sales 1993: 161)

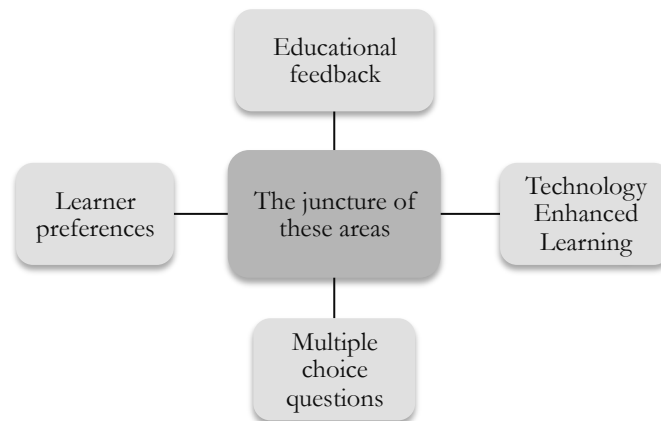
In this research TBI feedback is as defined as per (Mory 2003) but restricts the usage of the term to the context of technology based instruction. TBI feedback is defined as,

Any communication or procedure given to inform a learner of the accuracy of a response in the course of technology based instruction.

1.3 IMPORTANCE OF TOPIC

This research lies at the juncture of four areas of contemporary educational research: educational feedback, technology enhanced learning, multiple choice questions and learner preferences. In this section the importance of each area is introduced and, following this, the importance of focusing on the juncture of these areas is highlighted. This approach is illustrated in the following diagram.

Figure 1: The location of topic within educational research



1.3.1 Educational feedback

While the focus of this research is specifically TBI feedback, this section first discusses the importance of the more general concept of educational feedback. Consideration of this broader type of feedback allows conclusions to draw upon a wider body of research. All conclusions drawn are applicable to the specific area of TBI feedback.

The provision of educational feedback affects the efficiency and accuracy of learning (Bloom 1984; Clarina, Ross et al. 1991; Pashler, Cepeda et al. 2005) and is one of the 5 to 10 most significant positive influences on student achievement (Hattie and Timperley 2007: 82). Conversely, an absence of educational feedback can impact negatively on student achievement (Trowbridge and Cason 1932 cited in Bangert-Drowns, Kulik et al. 1991: 213; Laurillard 2002: 55) and lead to student dissatisfaction (Pargetter, McInnis et al. 1998; Boehler, Rogers et al. 2006).

Research into the role of educational feedback in the modern era can be found dating back to the beginning of the 20th century. An example is an experiment by Trowbridge and Cason (1932 cited in Bangert-Drowns, Kulik et al. 1991: 213) who asked blindfolded subjects to draw a series of lines exactly three inches in length. Subjects were given 100 attempts to perfect their efforts. Half the subjects were given feedback regarding the accuracy of their attempts and showed significant progress as they progressed through the trial. The remaining subjects received no such feedback and, as was expected, showed no improvement.

Learning is clearly possible without feedback. Taking observational learning (Bangert-Drowns, Kulik et al. 1991) as an example, knowledge or skills can be acquired from passive media such as books and television. Feedback

impacts on learning when we apply those knowledge and skills in an interaction with the world around us. As Laurillard (2002: 55) states,

“Action without feedback is completely unproductive for a learner.” (Laurillard 2002: 55)

Feedback serves to confirm the accuracy of existing knowledge, to confirm the efficacy of existing skills or to contradict either of these thus providing the opportunity to learn.

The importance of educational feedback to learning is affirmed by the central role played by feedback in much educational theory. Feedback forms a central role in educational paradigms such as behaviourism (Skinner 1938), experiential learning (Kolb and Fry 1975) and Laurillard's conversational framework (Laurillard 2002). Feedback also comprises one of Robert Gagne's nine events of instruction (Gagne 1985).

Empirical studies generally demonstrate that the provision of feedback, or lack thereof, can have a direct effect on the efficiency and accuracy of learning (Bloom 1984; Clarina, Ross et al. 1991; Pashler, Cepeda et al. 2005). Hattie and Timperley (2007: 82) refer to a synthesis, conducted by Hattie, of more than 500 meta-analyses of studies aiming to identify influences on educational achievement. This synthesis gathered data from 180,000 studies involving 20-30 million students and analysed the effects of more than 100 influences including attributes relating to the host institution, learners' backgrounds, the teacher and the curricula. Hattie's synthesis found educational feedback to be one of the 5 to 10 most significant influences on student achievement. More significant, for example, than socioeconomic influences and class size.

It has also been shown that learners themselves are aware of the importance of feedback in their studies (Pargetter, McInnis et al. 1998 ; Boehler, Rogers et al. 2006; Weaver 2006). In a 1998 study aiming "to identify and develop measures contributing to successful secondary to tertiary transitions", Pargetter, McInnis et al. (1998) found a perceived lack of feedback to be a primary cause of undergraduate student dissatisfaction, particularly among first year students. These findings are corroborated by Boehler, Rogers et al. (2006) who found feedback to be a key determinant of student satisfaction. Reflecting the results of these studies, feedback now constitutes one of the core items tested in the National Student Survey conducted each year in the United Kingdom by the Higher Education Funding Council for England (HEFCE 2010).

Despite this knowledge, there is growing evidence of dissatisfaction on the part of learners regarding the feedback they are provided. In the afore mentioned National Student Survey, the category 'assessment and feedback' is consistently the item obtaining the lowest levels of satisfaction across the survey (HEFCE 2010). In a review of contemporary literature, Huxham (2007) concludes,

“The quality of the feedback that students receive on assessed work remains a widespread source of dissatisfaction, and may even be declining.” (Huxham 2007: 601-2)

These studies indicate a requirement to improve the educational feedback provided to students and, as the review of the literature in this thesis proposes, a corollary requirement to conduct research into the nature of the feedback that should be provided.

1.3.2 Technology Enhanced Learning

The context of this research is technology based instruction which is defined here as being instruction delivered by a computer. Technology based instruction forms one component of the wider field of 'technology enhanced learning (TEL)' which can be defined as the use of information and communications technologies (ICT) in learning and teaching. TEL is also known as 'e-learning', 'learning technology' or 'educational technology'.

TEL can enable a more efficient use of resources within higher education, can widen participation and can increase student achievement (JISC 2008: 4). An attractive feature of TEL for educators is its ability to provide learners with educational feedback (Roper 1977 citing Gagne, 1974: 43; Laurillard 2002).

In 2008 the Joint Information Systems Committee (JISC) in the United Kingdom set out to identify what, if any, tangible benefits had arisen as a direct result of the higher education sector's investment in TEL. The subsequent report lists a range of distinct benefits and concludes that,

“The appropriate use of technology is leading to significant improvements in learning and teaching across the sector and this is translating into improved satisfaction, retention and achievement.” (JISC 2008: 4).

This broad statement reflects the view that TEL can make a positive contribution across many of the different components that form contemporary learning environments within higher education. For example TEL can enable a more efficient use of resources, can widen participation in higher education and can increase student achievement (Ibid. 2008: 2).

Diana Laurillard's 'conversational framework' model (Laurillard 2002) provides a useful framework when considering the question of how TEL can be used to promote learning. Laurillard proposes that dialogues, or conversations, between the different agents in a learning context are the basic requirement in any learning situation. Agents are defined as the student, the teacher and the environment. Using this model as a framework, the contribution of technology to learning can be seen as a facilitator of dialogue. For example, TBI promotes a dialogue between the learner and the learning materials through the provision of adaptable instruction and educational feedback; features not easily replicated using traditional learning materials such as textbooks. Communication technologies such as social networks, VOIP, instant messaging systems, discussion boards and wikis provide additional possibilities for dialogue between a teacher and a learner(s) and also between learners themselves.

That the ability of TEL to provide educational feedback to learners is its key advantage is supported by the educational psychologist Robert Gagne who states,

“The provision of immediate feedback in computer assisted instruction (CAI) is perhaps its most instructionally appealing feature” (Roper 1977 citing Gagne, 1974: 43).

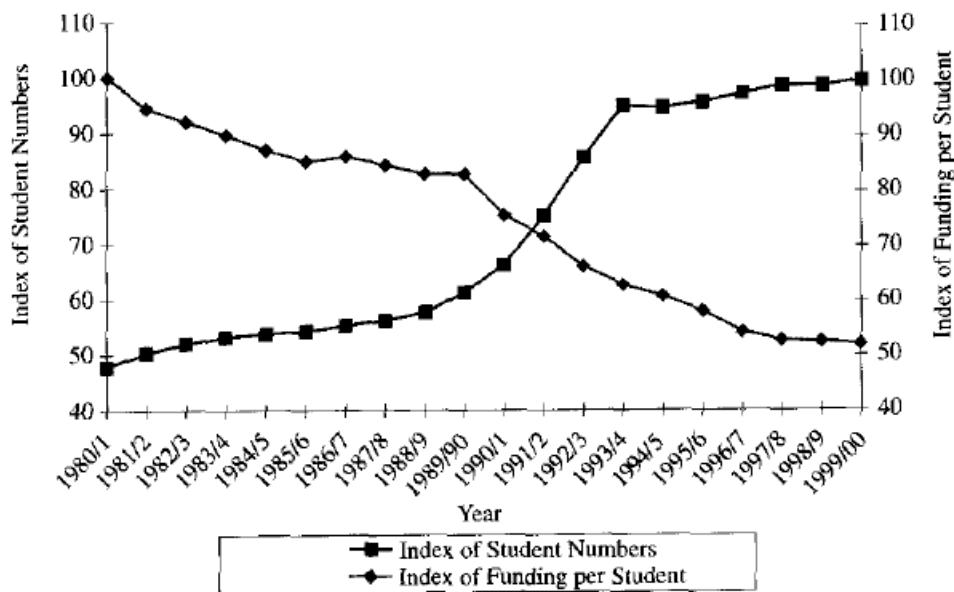
This ability is particularly significant in contemporary higher education contexts in which opportunities for teacher-to-learner dialogue are reduced by ever increasing class sizes. This is explored further in the following section.

1.3.2.1 *The expansion of education in the United Kingdom*

In 1962, the ‘Robbins Report’ commissioned by the British government noted that there were 118,000 students enrolled in universities in the United Kingdom and called for this number to increase to 560,000 by 1980 (Robbins 1963). This led to dramatic expansion of higher education in the United Kingdom. By 1971 there were 176,000 students in 45 universities (Farwell 2002: 151), and by 2002 there were 1.8 million students at more than 150 universities (Ibid. 202: 151). In the academic year 2009 to 2010 there were 2.6 million students studying at 165 universities (HESA 2011).

However, this expansion has not been accompanied by matched increases in funding. As figure 2 below taken from Greenaway and Haynes (2003) illustrates, real funding per student almost halved between 1980 and 1999.

Figure 2: Index of Student Numbers and Public Funding for Higher Education
(Greenaway and Haynes 2003: 152)



The ratio of number of students to number of staff increased during this expansion moving from an average 9:1 in 1980 to 17:1 in 1999 (Ibid. 2003: 154). In 2003 this ratio was 21:1 (AUT 2005: 2). The result has been larger class sizes (Greenaway and Haynes 2003: 153) and this has been shown to have a negative impact on student achievement Gibbs (2010). In a 2010 report into educational quality, Gibbs found that,

“Meta-analysis of large numbers of studies of class-size effects has shown that the more students there are in a class, the lower the level of student achievement” (Gibbs 2010: 19)

A link between large class sizes and declining student performance had been identified previously by Benjamin Bloom in a seminal paper (Bloom 1984). Bloom compared the performance of students under three different teaching conditions. 'Conventional teaching' in which students study in groups of 30 with one teacher, 'mastery learning' which differs from conventional teaching in that regular formative tests are provided then followed by corrective procedures, and 'tutoring' which differs from mastery learning in that the class size is reduced from 30 to between 1 and 3. Bloom found that the average performance of learners studying via tutoring was two standard deviations above the average performance of learners under the conventional teaching condition and one standard deviation above the average score of learners under the mastery learning condition. Bloom's study suggests that both class size and the provision of educational feedback affect learner performance. In fact, Bloom proposes that these two factors are inseparable in that it is the lack of educational feedback and an associated corrective process that constitutes the significant disadvantage of large class sizes.

As class sizes increase, the provision of educational feedback becomes increasingly challenging and Bloom is not alone in identifying this as the core issue. In a critique of contemporary teaching practice within higher education, Laurillard (2002: 81) proposed an insufficient provision of feedback on student actions to be the 'weakest link' in the established format of study.

Given the above, the most significant contribution made by Technology Enhanced Learning to Higher Education may be to provide educational feedback to learners in an environment where classes continually increase in size.

1.3.3 Multiple choice questions

During the studies that comprise this research, learners are presented with TBI feedback following their response to multiple choice questions. When compared to assessment questions graded by a human educator, the use of multiple choice questions can lead to an increase in the accuracy and objectivity of grading, to feedback being provided in a more timely manner and to the assessment of a wider range of course topics (Simkin and Kuechler 2005: 75).

Multiple choice questions are widely used. As Williams (2006: 288) observes, countries around the world, particularly the United States, have adopted multiple choice questions as the foundation of their testing systems. On economics courses in higher education it has been estimated that between 45 per cent and 67 per cent of assessment is via multiple choice questions (Simkin and Kuechler 2005: 74). Multiple choice questions also form a core component of exams leading to professional accounting qualifications such as those accredited by the Association of Chartered Certified Accountants (ACCA) in the United Kingdom and the American Institute of Certified Public Accountants (AICPA) in the United States.

The effectiveness of multiple choice questions as a form of assessment is a topic of ongoing debate. This section will summarise both sides of the argument but first the location of multiple choice question within the different categories of assessment question types is described.

Multiple choice questions fall under the category of ‘objective questions’ which can be defined as those which ‘require a user to choose or provide a response to a question whose correct answer is predetermined’ (CAA 2002). Objective questions are often discussed in contrast to ‘constructed response’ or CR type questions such as an essay topic or a poster presentation in which the learner determines, and constructs, the answer. Objective questions can take a variety of forms in addition to multiple choice. For example, matching exercises in which learners match an input item with a relevant target item, gap fill exercises in which students enter an appropriate response into a gap in a text and sorting activities in which learners are asked to sort a list of items into a specified order. However all these formats are, at root, a variation of the multiple choice question format and it is proposed that the findings of this research are equally applicable to these other forms of objective questions.

In a review of contemporary literature, Simkin and Kuechler (2005: 75) identify 16 advantages afforded by multiple choice questions, the majority of these relate to the fact that objective tests can be graded by a machine. These include increased accuracy and perceived objectivity in the grading process, feedback being provided in a timely manner and the ability for educators to ask questions on a wide range of topics.

Despite these advantages, multiple choice questions have been the target of much criticism (Brown, Bull et al. 1997; Bradbard, Parker et al. 2004; Buckles and Siegfried 2006). The primary focus of this criticism is that multiple choice questions do not test knowledge and skills to the extent that constructed response type questions are able. Research into such concerns has focused on the extent to which multiple choice questions test items in the Taxonomy of Educational Objectives identified by Bloom, Engelhart et al. (1956). The findings of Buckles and Siegfried (2006) articulate the argument as follows,

“We do not believe that Bloom’s highest levels of understanding, synthesis and evaluation can be accurately measured with multiple choice.” (Buckles and Siegfried 2006: 56)

However, research seeking to empirically test such concerns find that multiple choice questions suffer from no such shortcomings and are able to test the full range of educational objectives (Simkin and Kuechler 2005: 79). Simkin and Kuechler propose prior studies that resulted in conclusions to the contrary failed to take into account that some multiple choice questions aim to test lower level learning objectives and others higher. A further explanation behind inconsistencies in research findings may be the inherent difficulty in identifying and testing levels of student understanding across the taxonomy in a precise manner.

An additional focus of criticism concerning multiple choice questions is that a learner’s partial knowledge is not identified (Bradbard, Parker et al. 2004: 11). As Simkin and Kuechler state,

“(Multiple choice questions) deny a student the opportunity to organize, synthesize, or argue coherently, to express knowledge in personal terms, or to demonstrate creativity.” (Simkin and Kuechler 2005: 76)

Finally, multiple choice questions have been criticised from a practical perspective by highlighting the difficulty teachers face in their creation (Brown, Bull et al. 1997: 77).

Despite the controversy regarding the use of multiple choice questions, their use remains common within education and as such remain an important subject within educational research.

1.3.4 Learner preferences

Educational research can be considered as focusing on one or more of three agents within a learning context; the teacher, the learner(s) or the environment. The latter includes research focusing on the time and location of a learning activity or the use of materials such as simulations and multiple choice questions.

Research within the behaviourist paradigm in the early and mid 20th century investigated how interventions by the teacher and the manipulation of the learning environment could induce changes in a learner's behaviour. The learner's mind was treated as a 'black box' whose functions were not considered (Driscoll 2000: 33). The behaviourist paradigm has its origins B.F. Skinner's notion of operant conditioning and retains core concepts such as the reinforcement or punishment of behaviour (Skinner 1938). Inconsistencies in research findings and an apparent inability for behaviourist theory to account for complex learning tasks such as acquiring language led researchers to consider the role played by the learner's mind (Driscoll 2000: 66).

As discussed in the section 'The Constructivist perspective' (See page 30) in the literature review, constructivism has since emerged as the dominant paradigm in educational research. Constructivism places a primary focus on the learners who are considered, not as passive participants, but as the primary active agents in the learning process. A corollary of this view is recognition that the learners themselves ultimately determine what learning is acquired from an educational experience. As Hannafin (1992) put it,

“(The learner is) the principal arbiter in making judgments as to what, when, and how learning will occur.” (Hannafin 1992: 54)

Reflecting these standpoints, the constructivist paradigm has led to an increase in research focusing on the processes learners adopt when constructing knowledge. An investigation into the format for TBI feedback learners choose during TBI represents an example of such research.

1.3.5 The juncture of these areas.

During the previous sections, supporting arguments for the importance of feedback, technology enhanced learning, multiple choice questions and learner preferences were proposed. The importance of research focusing on the juncture of these areas is now argued.

Multiple choice questions survive the transition from the traditional learning environment to the technology enhanced learning environment as the arguments presented above supporting the use of multiple choice questions are equally applicable to the technology enhanced learning context.

It can be hypothesised that the adoption of multiple choice questions will continue to be pervasive. The arguments presented above for the importance of educational feedback are equally applicable within the context of technology enhanced learning and multiple choice questions are therefore a valuable focus for research. Furthermore, as argued above in the section ‘Technology Enhanced Learning’ (See page 19), the ability for multiple choice question based tests to provide educational feedback to learners in an environment in which classes continually increase in size, may become increasingly significant.

In addition, the capacity of higher education institutions to deliver multiple choice questions is increasing. All UK universities have adopted a virtual learning environment or VLE (Browne, Hewitt et al. 2010: 17) and all commonly adopted VLEs include the functionality to deliver multiple choice question based tests (JISC-Infonet 2009). Specialist testing systems, also capable of delivering multiple choice questions, are also common. One such software is used by approximately 2000 institutions in 50 countries (Respondus 2011).

While the use of multiple choice questions within higher education and may increase, prior research is inconclusive as to the relative effectiveness of different TBI feedback formats and the nature of TBI feedback that should be provided to learners is unclear (See section ‘Prior research on the taxonomy of TBI feedback types’ in the literature review (See page 46).

An improved understanding of learners’ engagement with feedback may help address this issue and there have been calls for research which places greater focus on the learner (Mory 2003: 771, Handley, Price, et al. 2011: 548). This research contributes to this call by investigating learners’ preferences for the format of TBI feedback provided during technology based instruction. This research asks the question, what format of TBI feedback do learners elect to view when given the choice?

1.4 RESEARCH QUESTIONS

The approach taken was exploratory in design however the course of this research was guided by the following two questions.

1. What format of technology based instructional (TBI) feedback do learners elect to view if given a choice?
2. What are the factors that affect this choice?

1.5 SUMMARY OF RESEARCH METHOD

A mixed-method approach was adopted consisting of four studies. Studies one, three and four comprised technology based observations in which a software tool, developed as part of this research, tracked subjects’ choices

relating to the format of TBI feedback while studying technology based instruction. Study two comprised a series of semi-structured interviews and investigated why subjects made their choices.

Over a three year period data were collected at one host institution relating to 24 cohorts of students studying one or more of three items of technology based instruction each containing between 15 and 50 multiple choice questions. In total, students responded to 61,929 questions.

This research first defined 'format' in relation to a taxonomy of TBI feedback formed following an analysis of prior research. This taxonomy consists of six feedback types, a measure of feedback 'load' and two conditions regarding the timing at which feedback is provided. The format of TBI feedback refers to a particular selection of attributes from this taxonomy.

Prior research was then reviewed in order to form hypotheses regarding the frequency at which learners will elect to view the different attributes in this taxonomy and also which factors potentially affect whether learners choose to view each feedback type.

Three items of technology based instruction were selected as the materials to be used and the feedback provided within these materials was adapted such that all attributes in the taxonomy of TBI feedback were represented. A software tool was then developed which delivered these materials to the research subjects, tracked their subsequent interactions in the software interface and provided the data gathered in an accessible format.

In studies one, three and four, subjects studied technology based instruction containing multiple choice questions and, following each question, were presented with six feedback options corresponding to the six types in the taxonomy of TBI feedback. Subjects were also given the choice either to view feedback immediately following their response to a question or after having answered all of the questions. The software tool recorded the target and timing of each 'click' subjects made in the interface.

Study two comprised a series of semi-structured interviews and investigated why subjects made the choices observed in the technology based observation that comprised study one.

1.6 CONTRIBUTION TO KNOWLEDGE

1.6.1 Contribution to theory

Feedback has a powerful effect on learning (Hattie and Timperley 2007: 83) however the format of TBI feedback that should be provided to learners remains unclear (Kulhavy and Stock 1989, Kluger and DeNisi 1996, Mory 2003, Shute 2008). Prior research has focused on the relationship between the format of TBI feedback and scores achieved in subsequent performance tests. As will be discussed in the literature review, an inconsistent pattern of results led to calls for research placing a greater focus on the learner (Mory 2003: 771, Handley, Price, et al. 2011: 548).

This research makes an original contribution to the literature by identifying the format of TBI feedback that learners prefer. This preference is established by observing the frequency at which learners elect to view different types within a taxonomy of TBI feedback and the timing at which they elect to view them. The results are organised into eight themes which are then unified and integrated with prior theoretical models to produce a model of learner interaction with TBI feedback. This model identifies a process through which learners interact with different TBI feedback attributes.

These results have implications for existing theoretical models. For example, the proposed existence of a 'feedback threshold' (Phye 1979, Kulhavy, White et al. 1985, Dempsey and Litchfield 1993) likely depends on the nature of the TBI and variation in the length of time learners' spend on feedback in Kulhavy and Stock's (1989) model of 'Response Certitude' is explained by variation in the feedback attributes towards which subjects give attention. This research also makes a novel contribution to the on-going debate regarding the timing at which TBI feedback should be provided. These implications are discussed in more detail in the section "Implications of the results" (See page 265).

In addition to furthering our understanding of how learners engage with TBI feedback, the results of this research reveal avenues for future research that may address the inconsistent pattern of results discussed above. This model contends that factors at the question level, learner level and environmental level affect learners' interactions with TBI feedback. This suggests that a lack of control for these factors has acted to confound prior results. This is discussed in more detail in the section 'Conclusions: Suggestions for future research' (See page 272).

1.6.2 Contribution to practice

The problem faced by teaching practitioners has been articulated as follows,

"Arguments over the relative merits of simple or complex feedback continue to be a problem for practitioners, who want a simple answer to what is a knotty question." (Dempsey, Driscoll et al. 1993: 25)

Despite making no claims as to the effectiveness of different feedback formats, the results of this research do make seven recommendations for good practice regarding the format and presentation of TBI feedback. These are provided as an appendix to this thesis, "Appendix two: Recommendations for practitioners".

In addition, the results of this research enable researchers and practitioners identify differences between the behaviour that learners prefer to adopt and any behaviours found to lead to more effective learning. As Laurillard (2002) states,

"An insight into the student's view of the learning process would give us some basis for deciding on a teaching strategy." (Laurillard 2002: 41)

2. LITERATURE REVIEW

This chapter reviews prior research relevant to the field of TBI feedback and starts with a description of the approach taken.

A discussion then follows on how research investigating TBI feedback has been approached from the behaviourist, cognitivist and constructivist perspectives. This is followed by an outline of five models of instructional feedback that potentially act as theoretical frameworks through which the results of the present research could be interpreted.

A taxonomy of TBI feedback types is then proposed following an analysis of how different formats of instructional and TBI feedback have been categorised within prior research. Research investigating the attributes within this taxonomy is then reviewed and the research gap is identified.

Finally, a further analysis of prior research is conducted in relation to the taxonomy of TBI feedback in order to form hypotheses regarding the format of TBI feedback learners prefer.

2.1 PROCEDURE

This literature review began by identifying the journals, databases and prior literature reviews that would be subject to a systematic search. This consisted of consultation with practicing researchers, reference to a previous PhD thesis into the field of TBI feedback (Handley 2003), exploratory searches of the EBSCO database and also of the online library at Imperial College London.

One output of this exploratory stage was a list of potentially relevant educational journals which was subsequently filtered using the criteria of quality and relevance. Relevance was determined by the degree of focus on either assessment or educational technology. The quality of journals was determined through consultation with mentors together with reference to the claimed impact factor. This process identified four journals, the British Journal of Educational Technology, ALT-J, Assessment & Evaluation in Higher Education and the Review of Educational Research to form the basis of a subsequent systematic search. Papers relating to TBI feedback within these journals were identified via a manual review of all articles appearing in each journal over the previous decade. This review was conducted in 2005 and resulted in 17 papers being identified for further study. This study resulted in four papers being determined as relevant to the present research. (Kulhavy 1977; Kulik and Kulik 1988; Bangert-Drowns, Kulik et al. 1991; Butler and Winne 1995).

The exploratory stage also identified two papers reviewing contemporary feedback literature (Handley 2003; Mory 2003). These two papers, together with the four articles identified during the journal search were then subject to the snowballing method. The bibliographies of these papers were analysed resulting in a larger number of papers for consideration. This same procedure was conducted on the additional papers and this iterative process

continued until potential papers for consideration either repeated themselves or did not meet the criteria of relevance and quality. In this manner, a list of 36 papers was identified for analysis.

In parallel to the snowballing method described above, I conducted systematic searches of the EBSCO database, the online library at Imperial College London and the online library at the Institute of Education, University of London. These searches adopted the inclusion criteria described below. These database searches resulted in a further nine papers being added to the list for analysis.

In addition, I conducted frequent, exploratory searches of the three databases throughout the period of this research in response to information gathered from peers or through additional reading. Similar searches were also conducted of the Google Scholar and Education Resources Information Center (ERIC) databases. These searches were less systematic in nature and were conducted to identify possible gaps in the more systematic searches or to pursue new avenues that arose as a result of the exploratory nature of this research.

In 2010, I repeated the systematic search of the four journals stated above and this resulted in a further nine articles being identified of which four were determined to be relevant. (Weaver 2006; Williams 2006; Hattie and Timperley 2007; Shute 2008). A subsequent snowballing method resulted in a list of 13 papers that were then subject to analysis.

These methods lead to a total of 58 papers that form the basis of this literature review. Of particular pertinence to this research are 15 papers that manipulated the format of TBI feedback in order to identify any effect on scores in subsequent performance tests. An analysis of these papers is included in appendix three to this thesis.

It should be noted that any literature search is restricted by the sources available to the researcher and while all papers deemed critical were obtained, there were instances when an article was eliminated from the review due to access restrictions.

Inclusion criteria

For the systematic database searches, the primary search terms were 'feedback', 'educational feedback', 'instructional feedback' and 'formative feedback'.

Subsequently, these same terms were also used in conjunction with terms relating to technology based instruction including 'technology enhanced learning', 'computer based learning', 'e-learning', 'educational technology' and 'learning technology'. And also in conjunction with terms relating to assessment such as 'assessment', 'multiple choice', 'objective questions' and 'educational assessment'.

No date criteria were applied when conducting a search.

Exclusion criteria

Only articles published in the English language were incorporated into this review.

2.2 PRESENTATION OF FINDINGS

2.2.1 Feedback as viewed through educational theory

As a subset of educational research, research investigating educational feedback such as TBI feedback tends to reflect and be underpinned by contemporary educational paradigms. Since the beginning of the last century three theoretical perspectives have dominated educational research; behaviourism, cognitivism and constructivism. The consideration of feedback under each of these perspectives is now discussed.

2.2.1.1 *The Behaviourist perspective*

Educational feedback received a great deal of attention in the early and middle to late 20th century as a key component of the behaviourist learning paradigm (Handley 2003), a perspective with strong links to Skinner's work on operant conditioning (Skinner 1938). In this model, the role of feedback is to act as a 'motivator' and a 'reinforcer' of correct behaviour, the learning and teaching equivalent of the food pellet used in operant conditioning (Mory 1992: 6). Thus studies on instruction focused on feedback as a 'reinforcer' and the role of feedback in correcting errors or providing information was largely ignored.

Many of the principal ideas that emerged during this time are reflected in Skinner's 'programmed instruction' model of teaching (Skinner 1968). In this model, learning is broken down into a series of short tasks. Learners progress from one task to the next upon submitting a correct response. The role of feedback is to reinforce correct responses to tasks. However, research largely failed to provide evidence for the effectiveness of the role feedback as a reinforcer. As Mory (2003) states,

“Around 1970, most researchers began to doubt the feedback as reinforcement view. In fact, 10 years of research under this paradigm showed no systematic effects for feedback. Studies provided little evidence that feedback following positive responses acts in a reinforcing manner (Anderson, Kulhavy, & Andre, 1972; Bardwell, 1981; Barringer & Gholson, 1979; Kulhavy, 1977; Roper, 1977).”

(Mory 2003: 746)

In addition, studies into feedback effects revealed phenomena not easily explained by the feedback as a reinforcer model. In a review of contemporary literature, Kulhavy and Stock (1989: 282) point to a number of papers from this period demonstrating one key such example, the delay-retention effect, in which a delay between the original instruction and the subsequent feedback leads to significantly better performance. This effect is not easily explained by the behaviourist model based on traditional operant-response experiments in which even a small delay should lead to a decline in learning and performance.

2.2.1.2 *The Cognitivist perspective (Information processing)*

The inability of the traditional behaviourist model to explain effects such as the delay-retention effect led researchers to explore models based on cognitive information processing approaches to educational psychology

(Driscoll 2000: 61). The primary theme of feedback research during the 1970s and 1980s was the role of feedback as an error correction mechanism. As Mory (2003) puts it,

“Examining feedback from an information-processing perspective, the learner participates in the system to correct his or her errors.” (Mory 2003: 747)

A typical study analyses errors made in an initial performance test and then investigates whether changes in the format of feedback provided result in the correction of such errors in subsequent tests. The wider availability of computers during this period also led to increased opportunities for technology based instruction and the beginnings of research into TBI feedback (Roper 1977; Kulhavy and Stock 1989).

However, while great progress was made in mapping the research landscape and in establishing that some feedback is better than no feedback, these studies often produced mixed or contradictory results. Reviewing the results of more than 50 such studies looking at the effects of changes in the format of feedback, Kulhavy and Stock (1989) conclude,

“We are unable to reach any useful conclusion regarding how the elaborative component of feedback operates on instructional responses.” (Kulhavy and Stock 1989: 289)

In addition, research from this period was unable to produce a theoretical model explaining the effects of TBI feedback that could be subject to empirical experiments.

2.2.1.3 *The Constructivist perspective*

Of the three educational paradigms considered here, constructivism is the dominant paradigm underlying contemporary educational research. To illustrate, a search of the Education Resources Information Center (ERIC) database for papers containing either of the terms “constructivism” or “constructivist” in their title or abstract and which were published between the years 1992 and 2012, yields 6,633 results. Equivalent searches for behaviourism and cognitivism yield 52 and 101 results respectively.

As Duffy and Cunningham (1996) note, the term constructivism has evolved to encompass a wide range of positions, all holding two tenets in common;

“Learning is an active process of constructing rather than acquiring knowledge, and
Instruction is a process of supporting that construction rather than communicating knowledge.”
(Duffy and Cunningham 1996: 171)

The origins of constructivism lie in the field of philosophy (Von Glasersfeld 1989: 122). To support this claim Von Glasersfeld points to the work of Giambattista Vico in the early 18th century who outlined a “thoroughly constructivist epistemology”. John Dewey, in the first half of the 20th century, was one of the earliest educational theorists to adopt constructivist ideas (Duffy and Cunningham 1996: 173). Dewey’s educational thought incorporates constructivist concepts such as the requirement for students to relate new learning to prior experience,

the importance of 'learning by doing' and a focus on the role of the teacher as a facilitator or guide as opposed to a didactic deliverer of knowledge. In the mid-20th Century, Jean Piaget, Ernst Von Glasersfeld and Jerome Bruner worked to build the foundations of contemporary constructivist thought.

When considering the range of educational theoretical views considered to be constructivist, Duffy and Cunningham (1996: 181) draw a distinction between the 'cognitive constructivist' and 'socio-constructivist' uses of the term. The cognitivist constructivist view emerged in the mid 20th century and is associated by Duffy and Cunningham with the work of Von Glasersfeld (1989) and Piaget (1967). Cognitive constructivism focuses on the individual and the strategies individuals adopt in order to make sense of their world. In the cognitivist constructivist view, new knowledge is obtained when learners encounter an experience that challenges their existing representation of the world. This is termed a perturbation. Learning occurs as learners resolve this perturbation and prior representations are refined to accommodate the new experience.

'Socio-constructivism' on the other hand has origins in the both the work of Vygotsky (1978) during the early part of the 20th century and in later ideas relating to 'situated cognition' (Lave and Wenger 1991) together with its related concept 'communities of practice'. Both Vygotsky and the work of Lave and Wenger emphasise the social aspect to knowledge formation stating that knowledge is constructed through interactions within a group context. Proponents of 'situated cognition' stress the importance of the situation in which learning occurs, proposing that knowledge has a contextualised character.

From both the 'cognitive constructivist' and 'socio-constructivist' viewpoints, the role of feedback is to instigate and then assist learners in resolving perturbations. Feedback can be considered to act as a guide, assisting learners in the process of acquiring knowledge during their engagement with instructional tasks. As Handley (2003) puts it,

"Education from a constructivist perspective is about 'drawing out' rather than 'putting in'. The purpose of feedback is to help learners make connections and explore their own understandings. Feedback according to the constructivist perspective moves away from evaluative judgements about 'right' and 'wrong' and is more about offering ideas and insights for further reflection."

(Handley 2003: 28)

Askew and Lodge (2000: 10) compare the role of feedback in the expert/learner scenario to that in a counsellor/patient scenario in that its primary role is to enable the learner to come to their own conclusions regarding the object of study.

In comparison to research conducted under the cognitivist paradigm, feedback has thus far received little attention from a constructivist perspective (Mory 2003:770). This is the larger gap in the literature to which this research seeks to contribute. Research of this nature will provide new insights and perspectives on the role instructional feedback plays in learning. As Mory (2003) states,

“Certainly the inconsistencies in the feedback literature warrant some fresh ideas and perspectives. This researcher proposes that feedback be critically examined within a paradigm that embraces the philosophy of constructivism.” (Mory 2003: 771)

In recent years, researchers have turned their attention to learners’ perspective of educational feedback (Weaver 2006; Handley and Cox 2007; King, Schrodtt et al. 2009; Walker 2009) and this is discussed in more detail in the section “The research gap” (See page 55) later in this literature review.

2.2.1.4 Note on learning styles

Related to constructivist thought is the development of a number of models of ‘learning styles’ each of which purports that learners are inclined to adopt one of a number of distinct approaches to learning. An early model is that of Kolb (1984). Kolb built on the works of John Dewey and Jean Piaget mentioned above to produce a theory of ‘experiential learning’ (Ibid: 4). Within this framework Kolb posits that learners tend to adopt one of four learning styles most strongly; these being a ‘converger’, a ‘diverger’, an ‘assimilator’ and an ‘accommodator’. Other models include that of Honey and Mumford (1992) and Entwistle’s “Approaches and Study Skills Inventory for Students” (Tait, Entwistle et al. 1998). These are three examples from a large number of models that have been proposed. One meta-study of research into the area of learning styles identified 71 such models (Coffield, Moseley et al. 2004).

Given the focus of the present research on learners and the preferences they hold for TBI feedback, the concept of learning styles is worthy of consideration. However, periodic reviews of the body of research within the field of learning styles point to significant contradictions in research findings and question the validity of the research on which models of learning styles are based (Reynolds 1997; Coffield, Moseley et al. 2004). Reynolds’ review, conducted within the context of management education, concluded,

“The prominence of learning style theory in management development is in spite of the considerable lack of research support for the concept.” (Reynolds 1997: 128)

In a broad and systematic review, Coffield, Moseley et al. (2004), questioned the validity of the research underlying all the models considered including those of Kolb, Honey and Mumford and Entwistle mentioned above stating,

“The sheer number of dichotomies betokens a serious failure of accumulated theoretical coherence and an absence of well-grounded findings, tested through replication.” (Coffield, Moseley et al. 2004: 137)

Given these questions regarding the validity of learning styles, such models are not included within the theoretical framework used to interpret the results of this research.

2.2.2 Models of learner interactions with instructional feedback

This literature review identified five studies that propose theoretical models of relevance to the present research. None of these models represents a contemporary, generalizable model of learners' interactions with TBI feedback however, despite this limitation, each contributes theoretical concepts valuable to the interpretation of results in the present research.

2.2.2.1 *Model 1: Kulhavy and Stock (1989) - The Place of Response Certitude*

In this research Kulhavy and Stock surveyed contemporary knowledge relating to TBI feedback and then adopted 'servocontrol theory' as a framework through which this knowledge was collated into a single model. Servocontrol theory emerged from the field of engineering but was applied by Kulhavy and Stock to the study of human learning.

Kulhavy and Stock's (1989: 281) model assumes that learners' interactions with a task occur across a three-stage cycle:

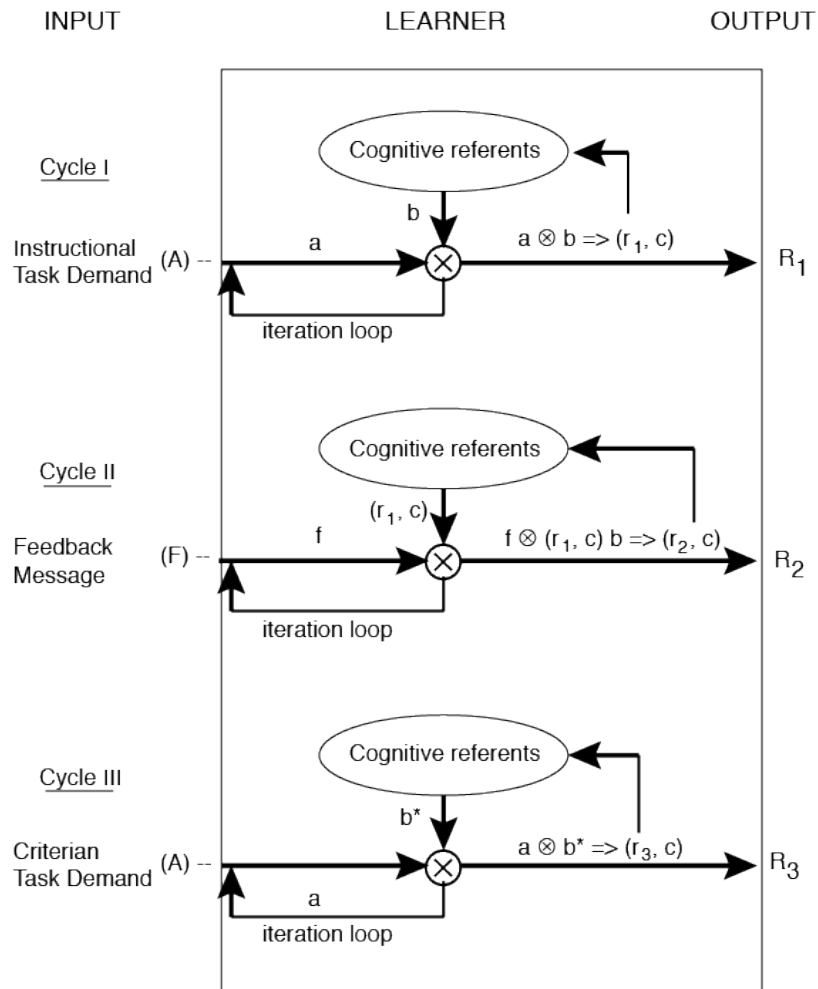
Stage one: The learner is presented with an instructional task and produces a response (R1)

Stage two: Feedback is presented, is processed by the learner to yield a behavioural response (R2)

Stage three: The learner is again presented with the task and produces a response (R3)

Terminology and processes from servocontrol theory are then applied to each of these stages to produce a model of learners' interactions with feedback. A detailed explanation of this model is beyond the scope of this review however figure 3 below illustrates its essence. 'Cognitive referents' are items of information stored within the brain to which a student refers when assessing the task. The symbol $\otimes$ represents a comparator that, for example, compares a feedback message to the cognitive referents.

Figure 3: The three-cycle feedback paradigm (Kulhavy and Stock 1989: 291)



The model is useful in that it describes a process through which learners engage with TBI feedback and an explanation for each stage of this process. However, this usefulness is tempered by the strong focus on one aspect of learners' engagement. As Kulhavy and Stock state,

“The hub of our theorizing is the variable of response certitude.” (Kulhavy and Stock 1989: 296)

Response certitude refers to the degree of certainty learners hold in their response and the result of focusing on this concept is that the predictions of the model are narrow. The model makes two key predictions. First, that response certitude affects the length of time learners spend processing a feedback message. Second, response certitude determines the degree to which feedback affects learners' performance in subsequent tests. The predictions are summarised in table 1 below.

Table 1: Key predictions from Kulhavy and Stock (1989)

	Low response certitude	High response certitude
Responds correctly	Low discrepancy. Shorter time required to process feedback. Lower probability that error will be corrected at stage three.	Low discrepancy. Shorter time required to process feedback. High likelihood that learner will make same correct response at stage three.
Responds incorrectly	As when responding correctly.	High discrepancy. Largest time required to process feedback. High probability that error will be corrected at stage three.

In addition the model predicts that, in the absence of feedback, the likelihood that a learner will respond to the same task with the same response increases as the degree of response certitude increases. The model also predicts that 'large scale' elaborative feedback will have the greatest effect when response certitude is low.

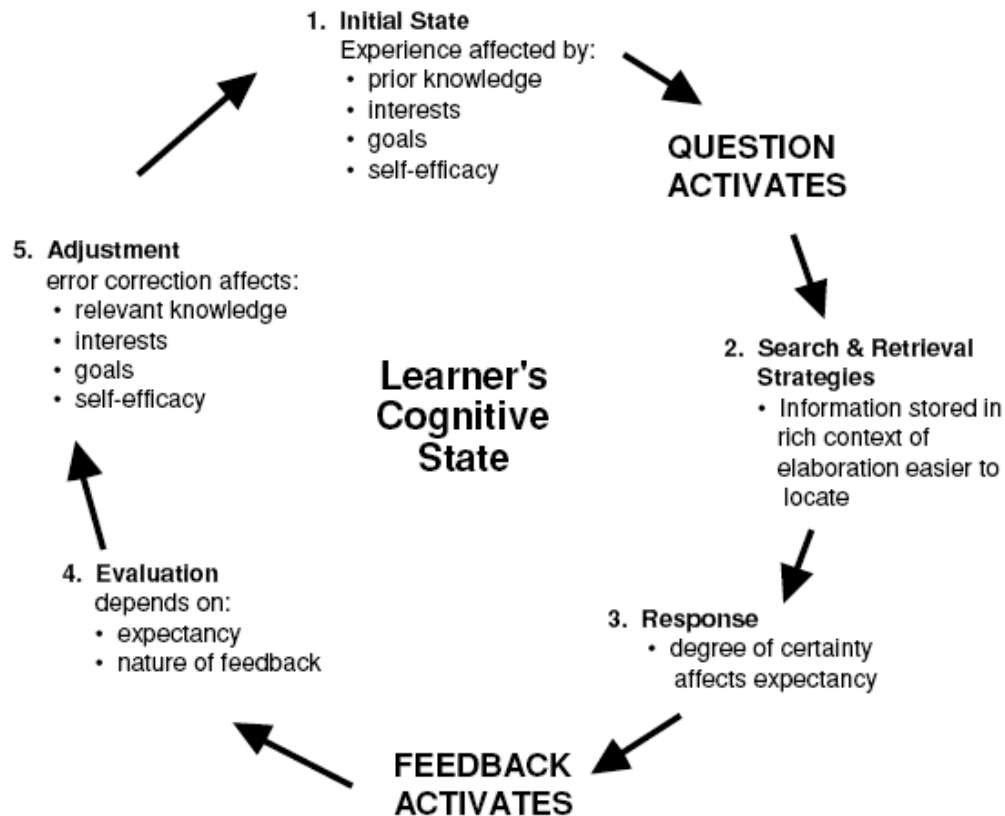
In the same paper, Kulhavy and Stock report on three experiments providing support for these predictions. As the authors note, these experiments all refer to a rather specialised learning design in which learners respond to a question, view feedback and then respond once more (Kulhavy and Stock 1989: 290). This may need to be taken into consideration when applying this support to learning designs in general

In addition to making rather limited predictions, an aspect of the model that reduces its relevance to the present research is that the theoretical framework through which the model was developed is strongly cognitivist in nature (Kulhavy and Stock 1989: 279). The model does not therefore draw on knowledge that has emerged in recent years from the constructivist perspective discussed in the previous section 'The Constructivist perspective' (See page 30).

2.2.2.2 *Model 2: Bangert-Drowns, Kulik et al. (1991) - A five-stage model describing the state of the learner receiving mediated intentional feedback*

Bangert-Drowns, Kulik et al. (1991) conducted a meta-study of 40 research papers in order to produce a model describing how learners interact with instructional feedback. As per Kulhavy and Stock's (1989) model, this aims to identify and describe the different stages learners pass through when engaging with instructional feedback. The result is the five-stage model as summarized in figure 4 below.

Figure 4: A five-stage model - Bangert-Drowns, Kulik et al.



(Dempsey, Driscoll et al. 1993 based on Bangert-Drowns et al., 1991: 40)

This model provides an alternative framework describing the process through which learners engage with feedback and has a simplicity lacking in the Kulhavy and Stock model described above. The model makes a valuable contribution by introducing the notion of ‘mindfulness’. A key conclusion of the underlying meta-analysis is that feedback can promote learning if it is received mindfully. This refers to the opposite of automatic, overlearned responses to feedback (Bangert-Drowns, Kulik et al. 1991: 217).

However, the usefulness of the model is compromised by a lack of detail relating to the nature of each stage. As Dempsey, Driscoll et al. (1993:) state,

“Although this model serves as an effective heuristic for organizing current findings regarding feedback, it is perhaps too general to do much more. In other words, its stages can hardly be argued. Without specific predictions, the model cannot be subject to empirical verification.” (Dempsey, Driscoll et al. 1993: 41)

As an example, stage four is the most relevant stage to the present study, however the findings of the Bangert-Drowns et al. meta-analysis add little insight to this stage beyond the findings of Kulhavy and Stock (1989), namely that the level of certainty held by learner, termed ‘response certitude’, will affect how long students spend on feedback following an incorrect response. Bangert-Drowns et al. themselves appear disappointed that their research was unable to shed more light on the learner state during this stage.

“Unfortunately, this meta-analysis was neither able to shed new light on the exact cognitive processes involved in response evaluation nor identify a characteristic of feedback that might influence this evaluation process.” (Bangert-Drowns, Kulik et al. 1991: 233)

Bangert-Drowns et al. do state that in addition to response certitude, a ‘genuine interest’ in the instructional content may be a determiner of engagement with feedback however this does seem to be based on common sense deduction rather than specific findings emerging from the meta-analysis.

As per Kulhavy and Stock, the theoretical framework adopted by Bangert-Drowns, Kulik et al. is cognitivist in nature. This is reflected in its focus on the mind as an information processing mechanism and the role of feedback as a corrector of errors (Bangert-Drowns, Kulik et al. 1991: 234).

2.2.2.3 Model 3: Butler and Winne (1995) – A model of self-regulated learning

Unlike the two models discussed above, Butler and Winne’s (1995) model is not one of instructional feedback but a model of ‘self-regulated learning’ that contains feedback as a key element. In the self-regulated learning model, focus is placed on the meta-tasks that learners perform during the study process. Such meta-tasks include setting goals, deciding upon study strategies and monitoring the outcomes of study behaviour. Not all learners are assumed to perform such meta-tasks effectively however proponents of self-regulated learning propose that those that do so will be more effective learners.

Within this model of self-regulated learning, feedback is a key component. Feedback is the device that allows learners to monitor their performance through a task. Butler and Winne’s model introduces the notion of internal feedback, that is feedback provided by a learner to his or herself. Internal feedback is used in contrast to external feedback which is that provided by an agent such as a computer, a teacher or a peer. The role of external feedback is described as follows,

“If external feedback is provided that additional information may confirm add to or conflict with the learner's interpretations of the task and the path of learning.” (Butler and Winne 1995: 248)

And if the latter then,

“Students may modify their engagement by setting new goals or adjusting extant ones; they may re-examine tactics and strategies and select more productive approaches adapt available skills, and sometimes even generate new procedures.” (Butler and Winne 1995: 248)

While not specifically a model of instructional feedback, the self-regulated learning aspect to Butler and Winne’s model introduces new considerations relating to feedback not found in previous models.

2.2.2.4 *Model 4: Handley (2003) - Learners' expectations of, and responses to TBI feedback*

In a series of studies, Handley investigated learners' expectations of and responses to TBI feedback. The key contribution of Handley's model is the introduction of the notion of 'closure'. Handley found the key expectation held by learners regarding TBI feedback to be as follows:

“Learners expected feedback to enable them to achieve a sense of closure to the task.”

(Handley 2003: 206)

With closure being defined as,

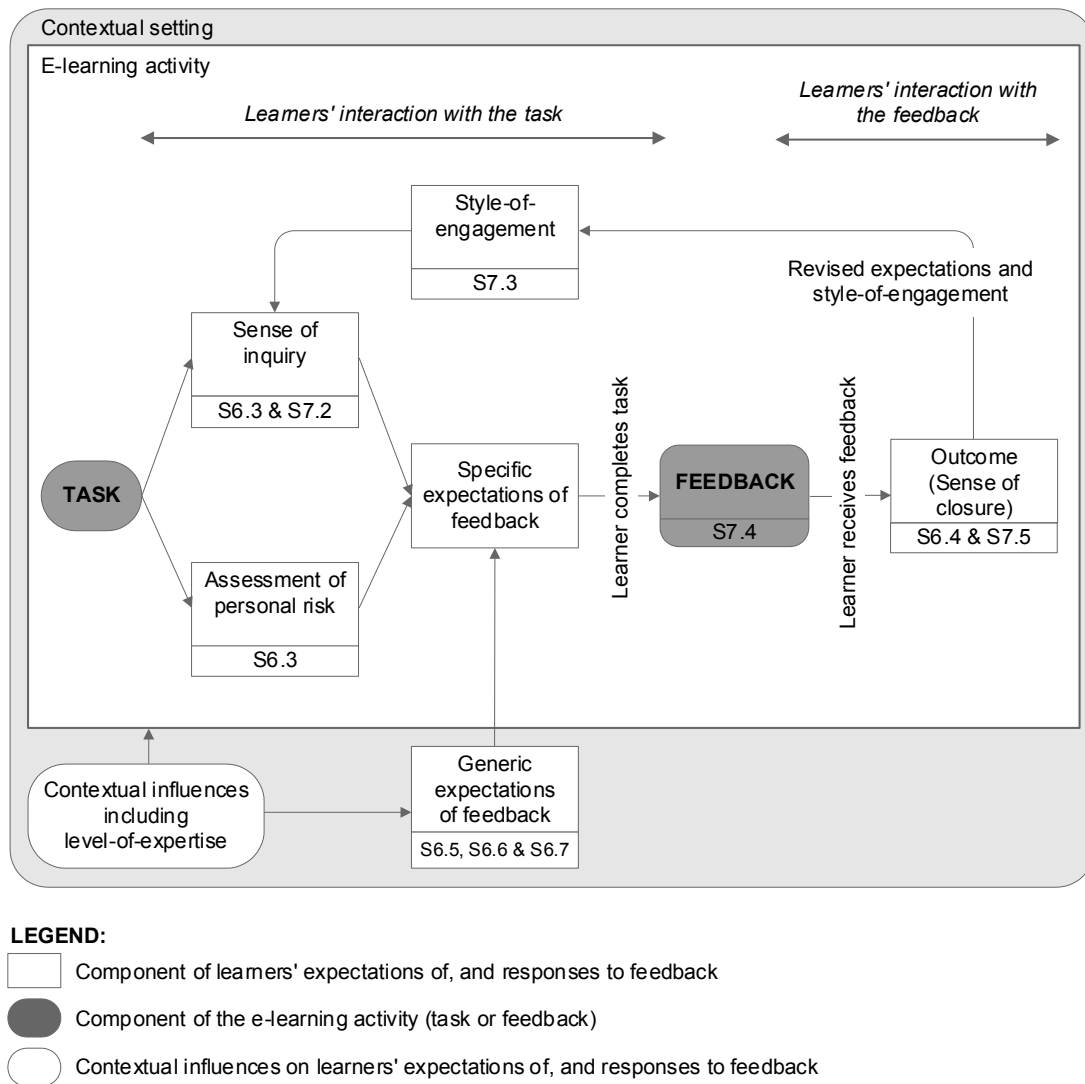
“The perception, on the part of the learner that expectations – arising specifically from the task, and generically from learners' expectations of what constitutes good feedback - are satisfactory met.”

(Handley 2003: 206)

Handley's studies focused on technology based training in the workplace, specifically the use of scenario based e-learning activities. The activities adopted focused on improving staff's problem solving skills in areas such as project management and ethics. Handley's central thesis is that learners seek closure from TBI feedback. Learners' achievement of closure is achieved firstly by satisfactory answering any questions they have relating to a task and secondly by meeting their expectations of feedback. These expectations are determined by four factors. First, learners' 'sense of inquiry' that is their degree of interest in the task arising from individual motivations and goals. Second, learners' assessment of personal risk which relates to an assessment of their ability to deal with a scenario in the real life context and the impact of failure. Third, learners' generic expectations regarding what constitutes 'good' feedback. Fourth, learners' 'style of engagement' with a task for which Handley specifies seven different types ranging from 'mindful learning' to 'going through the motions'.

Handley's model is summarized in the following diagram.

Figure 5: Handley's conceptual framework (Handley 2003: 206)



Of interest to educators is whether the achievement of closure depends on the nature of learners' engagement with an activity, their 'style of engagement'. For example whether learners engage with an activity in a mindful manner as generally desired by its designers or whether they merely 'play the game'. A number of researchers have proposed that learners tend towards the latter and choose the fastest route through 'TBI materials regardless of whether this route leads to effective learning of the material (Driscoll 2000: 389).

As described above, Handley's study found learners' 'style of engagement' to be a factor affecting the manner in which closure is achieved. A 'style of engagement' affects a 'sense of inquiry' which in turn gives rise to expectations relating to feedback that are required to be satisfied in order for closure to be achieved. For educators, a desirable outcome of Handley's study would be that the achievement of closure requires a mindful 'style of engagement'. However, this study found that this is not the case. Subjects adopting less mindful styles, for example "going through the motions" or "playing the game" were able to achieve closure on a task.

Handley's model introduces novel theoretical concepts such as the notion of closure which is adopted when interpreting the results of the present research. The key limitation of the model relates to generalizability. The

model provides a useful description of the behaviour of the learners observed throughout the four studies comprising the research. However the study took place within one organization and the TBI materials in each of the four studies were specific in nature being scenario based e-learning activities. As a result, not all the results are readily generalizable to the broader educational context. For example, one core component of the model, the 'impact of failure', does not apply to the materials adopted in the present research.

2.2.2.5 Model 5: Hattie and Timperley (2007) - A model of feedback to enhance learning

Hattie and Timperley's (2007) model does not focus on TBI feedback specifically but is a model of educational feedback in general. This model categorizes feedback into different types and levels before proceeding to state which of these types and levels are most effective.

A strength of this model is the size of the meta-analysis from which it is derived. The model draws on a synthesis of meta-analyses of studies focused on factors that influence educational achievement. The synthesis gathered data from more than 500 meta-analyses involving approximately 450,000 effect sizes from 180,000 studies, representing 20 to 30 million students. From this, Hattie and Timperley were able to establish the average effect size of feedback on performance to be 0.79 which constitutes one of the most significant effects.

"To place this average of 0.79 into perspective, it fell in the top 5 to 10 highest influences on achievement in Hattie's (1999) synthesis, along with direct instruction (0.93), reciprocal teaching (0.86), students' prior cognitive ability (0.71)" (Hattie and Timperley 2007: 83)

Hattie and Timperley then investigated the impact of different forms of feedback by focusing on 74 meta-analyses that included information relating to the form of feedback. Together, these meta-analyses contained data from 7,000 studies and 13,370 effect sizes. These included studies focusing on TBI feedback and, of interest here, is that Hattie and Timperley found TBI feedback to be one of the most effective forms of feedback.

"The most effective forms of feedback provide cues or reinforcement to learners; are in the form of video-, audio-, or computer-assisted instructional feedback; and/or relate to goals." (Hattie and Timperley 2007: 84)

From the 74 meta-analyses, Hattie and Timperley derive a model that first addresses the issue of what constitutes effective feedback. They propose effective feedback must answer three questions for both the learner and the provider of feedback.

Table 2: Questions to be addressed by effective feedback (Hattie and Timperley 2007: 86)

	Question asked by learner	Question asked by feedback agent	Term
1	Where am I going?	What are the goals?	'Feed up'
2	How am I going?	What progress is being made toward the goal?	'Feed back'
3	Where to next?	What activities need to be undertaken to make better progress?	'Feed forward'

The model then proposes that feedback should be directed at the appropriate level. The model identifies four different levels at which feedback can be applied. The first is the 'task level', feedback applied at this level relates to the outcome or product of a task and could include information relating to correctness or how performance could be improved. The second is the 'process level', feedback at this level focuses on the processes required to complete task. The third is the 'self-regulation level', this type of feedback encourages the learner to develop greater skill in conducting their own evaluation of performance in a task. The final level, the 'self' level, is feedback directed at the learner rather than the task. Feedback such as "You are a great student" is an example of feedback at the self level.

Hattie and Timperley conclude that the level at which feedback is directed affects its effectiveness. Feedback directed at the first level, the 'task level' is powerful when the information provided is useful for,

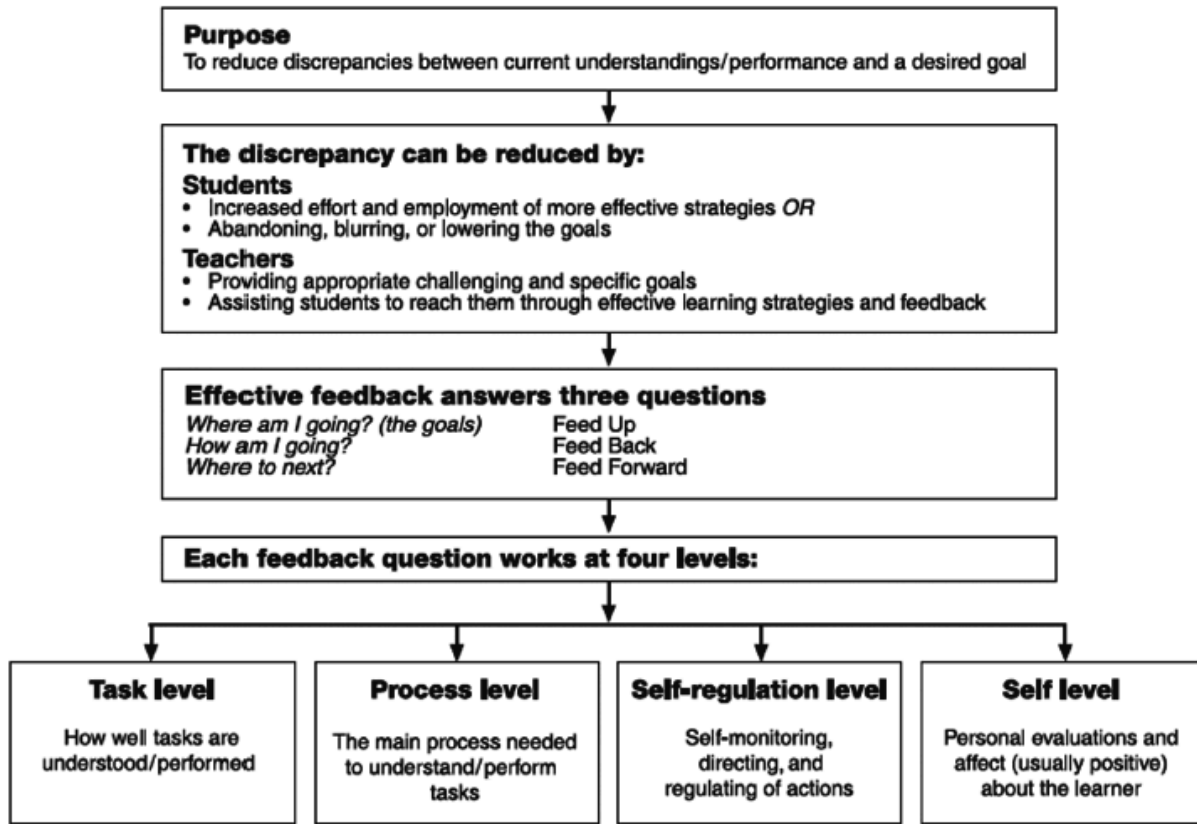
“...improving strategy processing or enhancing self-regulation (which it too rarely does)”.

(Hattie and Timperley 2007: 91)

Feedback at the next two levels, the 'process level' and the 'self-regulation' level are generally powerful. However, feedback at the final level, the self level, is the least effective level at which to direct feedback.

Hattie and Timperley's model of feedback is summarised in figure 6 below.

Figure 6: A model of Feedback to Enhance Learning (Hattie and Timperley 2007: 87)



In the context of the present research, a drawback of this model is that its utility is tempered by a focus on general educational feedback and not TBI feedback specifically. There is a strong focus throughout on feedback given from teachers to students in a classroom setting and the concepts of feed-up, feed-back and feed-forward operate largely at the course or the curriculum level and not the activity level at which TBI is generally utilized.

However, Hattie and Timperley’s model is comprehensive, supported by evidence and provides a further useful framework through which to consider the results of the present research.

This concludes the discussion of prior models of learner interactions with instructional feedback.

2.2.3 A taxonomy of TBI feedback

There have been a large number of studies on instructional and TBI feedback and in order to review these studies in a coherent manner this review now forms a taxonomy of six different TBI feedback types plus two timing conditions. Following this, research across the attribute of this taxonomy is then reviewed and discussed.

TBI feedback can take a number of different formats. An initial requirement in this research was to define what was meant by the term ‘format’. The approach taken was to survey prior research investigating the effect of TBI feedback format in order to produce a taxonomy of TBI feedback attributes. As Kulhavy and Stock state,

“We believe that the division of feedback into component structures more clearly directs attention to the specific functions performed by different elements.” (Kulhavy and Stock 1989: 303)

Initially collating the different formats of instructional feedback that have formed the foci of previous studies resulted in a taxonomy that was too broad for the purposes of a bounded study such as the present research. It was necessary to sharpen the focus and a narrower version of this taxonomy was developed by focusing on those categorizations of feedback of most relevance to the research topic. This process proceeded as follows.

At a high level, feedback can be classified as being either ‘intrinsic’ or ‘extrinsic’ or equivalently ‘task feedback’ or ‘external feedback’. Laurillard (2002: 55) defines intrinsic feedback as that “which is given as a natural consequence of the action” and extrinsic feedback as that “which does not occur within the situation but as an external comment on it”. An example of the former would be a student observing sales increase after increasing marketing expenditure in a business simulation. However, the present research focuses on the use of multiple choice questions and feedback in this context is primarily extrinsic. Thus intrinsic feedback was excluded from the taxonomy.

Feedback can also be classified as being either ‘adapted’ or ‘non- adapted’. Adapted feedback is that which changes in nature according to a learner’s response such that learners receive personalized feedback. However, feedback given in response to multiple choice questions is primarily non- adapted and therefore adapted feedback was also excluded from the taxonomy.

The literature review discovered a number of feedback formats that differ from those common to computer based testing systems. These included ‘try again feedback’ (also termed ‘answer until correct feedback’) in which the user is given more than one opportunity to identify the correct answer (Dempsey, Driscoll et al. 1993: 25), ‘error flagging’ in which the user is shown where errors exist in a response rather than being given the correct answer and ‘normative feedback’ in which the learner is shown how their performance compares to their peers (Kluger and DeNisi 1996). Studies focusing on ‘try again’, ‘error flagging’ and ‘normative feedback’ do exist however, reflecting the fact that they are not commonly used, this review found them to be few in number. As a result, these formats are not included in the taxonomy adopted in this research.

Thus, this research focuses on instructional feedback that is extrinsic and non- adapted. Attributes of this feedback are categorized here into three broad headings:

- The content of instructional feedback
- The timing of instructional feedback
- The medium of feedback

Prior research for each categories is now reviewed and discussed.

2.2.3.1 *The content of instructional feedback*

Kulhavy and Stock (1989: 285) review contemporary instructional feedback literature to produce the following categorization of for the content of feedback:

$$\text{Feedback} = \text{Verification} + \text{Elaboration (Type, Form, Load)}$$

The different components of this categorization can be explained as follows,

Verification

Verificational feedback informs the learner as to the correctness of their response.

Elaboration

Elaborative feedback is additional information that may guide the learner towards specified learning objectives. Elaborative feedback, also called extended feedback, can be considered in terms of type, form and load. Kulhavy and Stock specify three ‘types’ of elaborative feedback, which may be summarised as follows,

Table 3: Elaborative feedback types (Kulhavy and Stock 1989: 286)

Task specific	This feedback is drawn from the initial task, for example restating the correct answer or re-providing the incorrect answers.
Instruction based	This feedback consists of supplying relevant information from the course of study. For example the relevant part of the lecture notes, an explanation of why a response is incorrect or a restatement of the relevant rule.
Extra instructional	This is feedback consisting of additional information that has not been supplied to the learners previously.

The ‘form’ of elaborative feedback refers to changes in structure between the initial instruction and the feedback message while ‘load’ refers to the volume of information given during instructional feedback.

2.2.3.2 *The timing of instructional feedback*

In addition to the content of feedback, studies have considered the timing of feedback (Kulhavy 1977, Dempsey and Wager 1988, Morrison, Ross et al. 1995, Dihoff, Brosvic et al. 2003, Butler Karpicke et al. 2007), that is the moment at which feedback is delivered. In instructional feedback research, timing has generally been viewed as a dichotomous variable. Studies focus on whether differences arise between giving feedback immediately following a response and following a delay.

Dempsey and Wager (1988) suggest the following taxonomy for the timing of TBI feedback.

Table 4: A taxonomy for the timing of TBI feedback (Dempsey and Wager 1988)

Timing type	Description	Attributes
Immediate feedback	Is informative, corrective feedback given to a learner or examinee as quickly as the computer's hardware and software will allow during instruction or testing.	Item-by-item
		Learner controlled
		Logical content break
		End-of-module
		Break by learner
		Time controlled
Delayed feedback	Is informative, corrective feedback given to a learner or examinee after a specified programming delay interval during instruction or testing.	Item-by-item
		Logical content break
		Less than one hour
		1-24 hours
		1-7 days
		Extended delay (end-of-session)
		Before next lesson

2.2.3.3 *The medium of feedback*

This refers to the medium through which such content is delivered. For example whether a feedback message is delivered using words, pictures, audio, video or a combination of such media. A reasonable question would be to ask whether the medium of TBI feedback affects its effectiveness. Consider the feedback to a question testing a learner's knowledge of the orbitary movements of the earth, the moon and the sun. Using common sense, an educator may hypothesise that a video showing the moon orbiting the earth and both in turn orbiting the sun will be more effective in communicating this message using an equivalent text or audio description.

At present, there is a scarcity of research focusing on the media of feedback. However, there is an on-going and vigorous discussion on the relative effectiveness of different media within research on technology based instruction. The different positions are encapsulated in the debate between Richard Clark and Robert Kozma. Clark (1983) asserts that the medium of instruction is not significant, stating that the medium,

“Does not influence student achievement any more than the truck that delivers our groceries causes changes in our nutrition.” (Clark 1983: 445)

In contrast, Kozma (1991) asserts the following position,

“The research reviewed in this article suggests that capabilities of a particular medium, in conjunction with methods that take advantage of these, interact with and influence the ways learners represent and process information, and may result in more or different learning when one medium is used compared to another, for certain learners and tasks.” (Kozma 1991: 179)

This debate is unresolved. While Kozma’s position appeals to common sense, advocates of Clark’s position point to a lack of evidence supporting Kozma’s claims (Clark, Yates et al. 2010: 265). This research, takes the view that a consideration of the medium of feedback is outside the scope of this research. The medium of feedback is not considered.

2.2.3.4 *A taxonomy of TBI feedback attributes and associated values*

Following the above, this research adopts the taxonomy of TBI feedback types, load and timing conditions shown in table 5 below.

Table 5: A taxonomy for the timing of TBI feedback

TBI feedback attribute	Sub classification	Values
1. Content	Verificational	No feedback (NF) Knowledge of result (KR)
	Elaborative	Knowledge of correct response (KCR) Explanatory feedback (EX) Instruction based (IB) Extra-instructional (EI)
2. Load		Continuous variable
3. Timing		Immediate (I)
		Delayed (D)

This taxonomy includes all the attributes proposed by Kulhavy and Stock (1989) save ‘form’ which is considered too specific in that its use is restricted to a limited number of testing contexts. The taxonomy considers feedback consisting of explanations, included in Kulhavy and Stock’s ‘instruction based’ category, to be a separate classification termed ‘explanatory feedback’. In this taxonomy ‘instruction based feedback’ refers specifically to the re-presentation of course material. The taxonomy includes timing as a feedback attribute taking the values ‘immediate’ or ‘delayed’.

In this research, the ‘format’ of feedback refers to a particular selection of feedback attributes from this taxonomy.

2.2.4 **Prior research on the taxonomy of TBI feedback types**

The previous section surveyed prior categorizations of feedback to produce a taxonomy of feedback attributes under three headings; content, load and timing. Prior research relating to this taxonomy has investigated the impact of different formats on learners’ subsequent performance. Such studies first determine a classification of different instructional feedback types and then test whether changing the feedback type has a significant effect on learners’ performance in a subsequent test. Results have been inconclusive and frequently contradictory.

The inconsistent pattern of results has long been acknowledged (Kulhavy and Stock 1989, Kluger and DeNisi 1996, Mory 2003, Shute 2008) and following statement by Kulhavy and Stock in 1989 remains valid,

“Based on our assessment of the literature, we are unable to reach any useful conclusion regarding how the elaborative component of feedback operates on ... instructional responses.”

(Kulhavy and Stock 1989: 289)

Mory (2003: 755) reviewed research since Kulhavy and Stock’s statement and reached the same conclusion. As Shute (2008) observed, research since Mory has also been inconclusive,

“Researchers who have tackled the tough task of performing meta-analyses on the feedback data use descriptors such as “inconsistent,” “contradictory,” and “highly variable” to describe the body of feedback. Ten years later those descriptors still apply.”

(Shute 2008: 156)

The following sections summarise this prior research under each of the three primary categories in the taxonomy. Throughout, two possible causes of the inconsistent pattern of results are highlighted. First, a lack of consistency across the design of the studies and second, a lack of control for possible confounding factors. In the closing part of this section, these issues are discussed in more detail and additional explanations that have been proposed are discussed.

2.2.4.1 Prior research investigating the content of TBI feedback

The content of feedback has been the primary focus of prior research and a large number of studies have investigated whether the content of instructional feedback affects subsequent performance (Hanna 1976; Clarina, Ross et al. 1991; Morrison, Ross et al. 1995; Goodman 1998; Boehler, Rogers et al. 2006).

An example of such studies is Hanna (1976) who investigated post-test performance of 1,400 learners following MCQ based instruction during which the learners were provided with either no feedback, KR type verification feedback or KCR type elaborative feedback. Hanna found that the provision of either feedback type led to significantly higher performance. Another example is that of Roper (1977) who found differences in the effect of three TBI feedback types on learners’ subsequent performance in tests comprising part of a teaching programme in statistics.

“Feedback providing the correct answer was found to be superior to feedback simply saying whether the student's answer was correct or wrong; and this in turn was found to be superior to the total absence of feedback.” (Roper 1977: 43)

The results of these two studies are inconsistent. Roper found KCR type feedback to have a greater positive effect on subsequent performance than KR type feedback. However, Hanna found no significant difference. Such inconsistencies are frequent throughout prior research. Pridemore and Klein (1995) found a significant difference between the effects of KCR and EX type feedback while Gilman (1969) and Dempsey and Litchfield (1993) did

not. Moreno (2004) found subjects presented with IB type feedback scored significantly better in a subsequent performance test than those presented with KCR type feedback but this result is contradicted by the results of both Peeck (1979) and Andre and Thieman (1988).

It is notable that there is a great deal of variability in the design of these studies. Subjects range from junior high school students (Pridemore and Klein 1995), to high school students (Roper 1977, Birenbaum Tatsuoka 1987), to undergraduate students (Dempsey Litchfield 1993, Dihoff, Brosvic et al. 2003, Moreno 2004) and to subjects in an online research subject panel (Pashler, Cepeda et al. 2005). The subject matter varies from the qualitative (Kulhavy et al. 1985, Dihoff, Brosvic et al. 2003) to the quantitative (Roper 1977). Test questions sometimes focus on the simple recall of material (Pridemore and Klein 1995) and sometimes on the application of concepts (Birenbaum Tatsuoka 1987). Of potentially most significance, the combination of feedback types presented in prior research is rarely consistent. This literature review identified 15 prior studies investigating how changes in feedback type affected performance. A summary of these studies is presented in appendix three to this thesis. One combination of feedback types appears in three of these studies, two combinations appear in two studies each and the remaining eight combinations are unique to a particular study.

It may be that such variability in the research design is not significant however this has not been investigated. That factors such as prior knowledge, innate ability, age, teaching, class size, feedback, socioeconomic factors and affective factors affect educational performance is well established (Hattie and Timperley 2007: 83). It can therefore be hypothesised that some of these factors may also affect the impact of TBI feedback on performance. However, of the 15 studies presented in appendix three, only two controlled for such potential confounding factors. Pridemore and Klein (1995) controlled for the gender of the subjects and Moreno (2004) controlled for subjects' prior knowledge. The remaining studies make no reference to such controls.

Despite these inconsistencies, an analysis of prior research across the taxonomy of TBI feedback types does reveal broad themes on which there is some agreement.

The provision of some feedback is better than no feedback at all.

Of the 15 performance based studies presented in the appendix three to this thesis, 9 included NF (no feedback) as a feedback type and 7 of these 9 studies found that the provision of some feedback resulted in significantly higher scores in subsequent performance tests. In addition, there are plausible explanations why the remaining two studies did not find a significant difference. The first study, Gilman (1969) did find that the provision of KCR (Knowledge of correct result) and EX (Explanatory) type feedback resulted in significant differences in subsequent performance but failed to find a difference between KR (Knowledge of result) and no feedback. However, this is likely explained by the fact that subjects in this study were presented with questions repeatedly until all questions had been answered correctly. Thus the format of the test itself acted as a form of KR type feedback and it could be argued that the 'no feedback' condition was invalid.

The second study, Pridemore and Klein (1995), was based on an experiment investigating the effects of TBI feedback on learner performance and attitude. The testing instrument was an item of technology based instruction

containing presentational materials followed by multiple choice question based practice tests. The TBI feedback given in these practice tests formed the focus of research. The instruction aimed to teach learners how to use a microscope. Pridemore and Klein varied the type of control (allowing learners to choose whether to view the practice tests or making these mandatory) and level of TBI feedback (NF, KCR, EX) as independent variables and analysed the effect on achievement and attitude. Pridemore and Klein found that subjects receiving either NF or EX feedback performed significantly better in subsequent scores than those receiving KCR type feedback. In other words the provision of KCR type feedback was detrimental to subsequent performance when compared to the provision of no feedback at all. However, again there is a plausible explanation for this in the format of the study. Subjects were given access to all of the instructional materials while taking the practice tests. Pridemore and Klein suggest that those subjects receiving no feedback made efforts to establish whether they had responded correctly or incorrectly to a question by rereading the relevant section of the instructional material. Thus, as per Gilman (1969), it could be argued that subjects under the 'no feedback' condition were in fact receiving feedback.

It is concluded here that the provision of feedback, of any type in the taxonomy, has a more positive effect on performance in subsequent tests than not providing feedback.

The provision of elaborative feedback is more effective than verificational feedback alone.

Five of the 15 studies compared the effects of elaborative and verificational feedback on performance in subsequent tests. Four of these five tests found a significant positive performance benefit for the provision of elaborative feedback compared to verificational feedback alone (Gilman 1969; Roper 1977; Birenbaum and Tatsuoka 1987; Dempsey and Litchfield 1993; Pashler, Cepeda et al. 2005).

While the number of studies is small, this does suggest that the provision of elaborative feedback is more effective than that of verificational feedback alone.

Prior research is inconclusive as to the relative effectiveness of the different types of elaborative feedback

While providing elaborative feedback appears to improve learners' subsequent performance, the nature of the elaborative feedback that should be provided remains unclear. In a comprehensive review of contemporary literature, Kulhavy and Stock (1989) concluded,

“(The) feedback elaboration literature is comprised of opposing and inconsistent results.”

(Kulhavy and Stock 1989: 286)

Kulhavy and Stock (1989) identified more than 50 studies that manipulated the format of elaborative feedback to identify any changes in subsequent scores in performance tests and found no consistent pattern of results.

This literature review identified four studies focusing on the format of elaborative feedback published after Kulhavy and Stock's review (Andre and Thieman 1988; Dempsey and Litchfield 1993; Pridemore and Klein 1995; Moreno 2004). Two of these studies found that the format of elaborative feedback impacted on subsequent performance, two found that it did not.

Dempsey and Litchfield (1993) found no significant difference between the impact of different formats of elaborative feedback. This study was wide ranging in scope and will be referenced a number of times during this review. As such I will briefly explain this study in more detail. In this study 153 undergraduate biology students first studied a two-hour item of TBI focusing on substance abuse and chemical dependency. Multiple choice questions were presented throughout the material. The instruction was adaptive in that the materials presented depended to a degree on subjects' performance when answering the questions. After studying the TBI subjects sat a retention test comprising 24 questions testing the entire material. The independent variables in the study were the type of error made and the type of feedback provided. Subjects were randomly assigned to four different TBI feedback conditions:

- Knowledge of correct response

e.g. "That's it!" or "No. The correct answer is the binding of postsynaptic receptor sites"

- Knowledge of correct response and forced correct response

e.g. "No. The correct answer is the binding of postsynaptic receptor sites. Type the letter of the correct answer to continue"

- Knowledge of correct response and anticipated wrong answer remediation

e.g. "No. The correct answer is the binding of postsynaptic receptor sites + explanation of why other answers are incorrect"

- Knowledge of correct response and a second try

e.g. "No. The correct answer is (a) -- the binding of postsynaptic receptor sites"

The study focused on the possible effects of these independent variables on four dependent variables; performance in the retention test, feedback study time, on-task performance and feedback efficiency. Of interest here is the result that feedback type had no significant effect on subsequent performance in the retention test.

This result, that the type of elaborative feedback makes no difference to subsequent performance, is supported by Andre and Thieman (1988). However the remaining two studies, published after Kulhavy and Stock's review, found that the type of elaborative feedback did make a significant difference (Pridemore and Klein 1995; Moreno 2004).

These four studies are subject to the same issues of variability in research design and absence of experimental controls discussed above. It is also notable that the subjects in Moreno's study did not direct significant attention to the elaborative feedback provided. Moreno found that the group of subjects provided with elaborative feedback did not spend significantly more time on the activity than the group who were not (Moreno 2004: 105). One possible factor is that subjects' motivation may have been affected by undertaking the activity in experimental conditions and studying materials unrelated to their field of study.

However, these four studies show that the results of research since Kulhavy and Stock's (1989) review have been equally contradictory and the body of literature relating to the relative effectiveness of the different types of elaborative feedback remains inconsistent.

2.2.4.2 *Prior research investigating the load of feedback*

In comparison to the content of feedback, there has been little research investigating the effect of load of feedback. This review identified just two such studies.

Kulhavy, White et al. (1985) used the term 'feedback complexity' as a measure of feedback load. 120 subjects first read a text on the subject of naval strategy consisting of 16 paragraphs averaging 149 words in length. Following each paragraph subjects were presented with a multiple choice type question with five possible answers. Following a subject's response to the question, they were presented with feedback at one of the following four levels of complexity.

Figure 7: The four levels of complexity (Kulhavy, White et al. 1985: 287)

Level of complexity	Feedback components
I	a)
II	a) + b)
III	a) + b) + c)
IV	a) + b) + c) + d)

Where the feedback components are a) KCR type feedback comprising the presentation of the correct response, b) KCR type feedback comprising the presentation of the incorrect responses, c) EX type feedback comprising explanations as to why a response was incorrect and d) IB type feedback comprising the relevant section of the paragraph being tested.

Kulhavy, White et al. found that that the more complex feedback types did have a small positive effect on subsequent performance but questioned whether these effects justified the significant amount of student attention they required. In addition, the study found that the simplest level of feedback, level I, had the largest positive effect on subsequent performance. Two explanations for this result are offered. First, the EX type feedback comprising explanations to why incorrect responses are incorrect may actually serve to reinforce the error made and result in errors being repeated. Second, additional information beyond KCR may be perceived by learners are not being directly related to the error made and thus processed at a 'shallower' level.

The findings of Kulhavy, White et al. (1985) are corroborated to a degree by Phye (1979) who varied the amount of feedback provided to subjects and measured the subsequent effect on performance. The focus of Phye's study was narrow, focusing solely on providing KCR type feedback plus limited forms of EX type explanatory feedback.

Varying the load comprised presenting a subject with an increasing number of incorrect responses in addition to identifying the correct response. Some of these incorrect responses were not presented as options when a question was originally posed to the learner and as such the study does seem somewhat inauthentic and contrived. However, the study found that when more than sufficient information is provided in feedback then this additional information does not provide benefits in terms of increased scores in subsequent performance tests. Phye termed this proposal the ‘threshold effect’.

Taken together these two studies are inconclusive. Kulhavy, White et al. present a strong case for arguing that simpler versions of feedback are more instructionally efficient but data regarding the relative effects of the different levels of feedback on subsequent performance is incomplete and inconclusive.

Related to the issue of feedback load is ‘feedback efficiency’. Generally, elaborative feedback is more complex than verification feedback and requires a significantly higher investment of time and resources to develop. Such feedback also requires a significantly higher investment of time and effort by the learner. An important question for educators is whether any learning benefits justify such an investment. Kulhavy, White et al. (1985) explicitly address the issue of feedback efficiency and conclude that elaborative feedback beyond KCR type feedback is not efficient and does not represent a worthwhile investment of learners’ time. Thus supporting Phye’s (1979) suggestion of a ‘threshold effect’ at the level of KCR type feedback.

Dempsey and Litchfield (1993), described above, also attempt to determine the relative efficiencies of different feedback types adopting similar measure of efficiency to that developed by Kulhavy, White et al. (1985: 289).

$$\text{Feedback efficiency} = \frac{\text{correct retention percentage score}}{\text{sq root of feedback study time}}$$

(Dempsey and Litchfield 1993: 8)

As per Kulhavy, White et al. this study established that students spent more time studying the elaborate types of feedback but also found no difference in performance in the retention test across the four conditions. Using the measure of feedback efficiency above, they therefore concluded that KCR feedback is the most efficient type of feedback among the four.

2.2.4.3 *Prior research investigating the timing of feedback*

In 1977, Raymond Kulhavy supported the use of delayed feedback due to a phenomenon called the ‘delay-retention effect’ or the DRE (Kulhavy 1977: 214). The delay-retention effect was first identified by Brackbill and Kappy (1962) who found that delaying the provision of instructional feedback on a multiple choice test by a day or more improved learners’ scores in a subsequent retention test. Following Kulhavy’s paper, the effects of the delay-retention effect were corroborated by a number of further studies however, findings have not been unequivocal and, in a meta-study investigating possible reasons for this, Kulik and Kulik (1988) found that the delay-retention effect occurs only under certain, contrived test conditions and was not found when tests were performed under the conditions found in typical educational contexts. In a later review, Shute (2008) reaches the same conclusion.

“One interesting observation is that many *field* studies demonstrate the value of immediate feedback whereas many *laboratory* studies show positive effects of delayed feedback” (Shute 2008: 165)

Kulik and Kulik propose immediate feedback to be more effective than delayed feedback however Dempsey, Driscoll et al. (1993) go further suggesting that

“...delayed feedback, in many instructional contexts, is tantamount to withholding information from the learner”. (Dempsey, Driscoll et al. 1993: 24)

However, three subsequent studies revived the debate. Butler Karpicke et al. (2007) found that delayed feedback led to,

“...superior final test performance relative to immediate feedback”. (Butler Karpicke et al. 2007:273)

Morrison, Ross et al. (1995) also found that delayed feedback was more effective than immediate feedback but only for “lower level” learning. However, Dihoff, Brosvic et al. (2003) found that subjects receiving immediate feedback performed significantly better in a post performance test. It is not the case that a clear division can be drawn between these studies in terms of being either ‘field studies’ or ‘laboratory studies’. The literature remains inconclusive.

2.2.5 Explaining the inconsistencies in prior research on the taxonomy of TBI feedback types

As discussed in the previous section, prior research investigating the relative effectiveness of different formats of TBI feedback has been inconclusive. This possibly results from the variability across the design of these studies and the absence of controls for possible confounding factors. This section summarizes additional explanations that have been proposed and, in conclusion, questions the efficacy of the performance focused approach.

Kulhavy and Stock (1989) note that few studies have included an analysis of the nature of errors made by learners during instruction, and propose this to be a confounding factor.

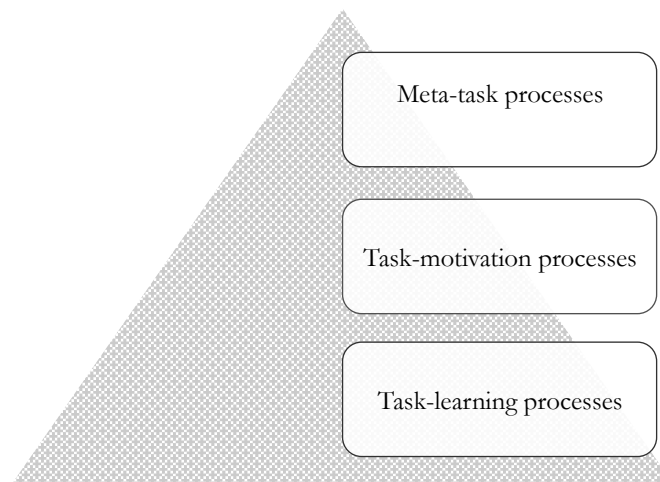
“Unfortunately, few of these studies manipulate (elaborative) in any systematic manner, and most analyse only overall post- test performance, as opposed to treating initial errors and corrects separately. There is little indication in this research of how various elaborations might work to perform the elimination, substitution, and strengthening functions that we perceive to be the essential purposes of feedback.” (Kulhavy and Stock 1989: 286)

A more comprehensive attempt to explain the inconsistencies in performance related research is that of Kluger and DeNisi (1996) who propose the primary issue to be the absence of a theoretical framework explaining the relationship between feedback and performance.

“Finally, we believe that the major culprit is a lack of a general theory regarding the effects of FI (Feedback Intervention) on performance. Without a comprehensive theory, there is no way to integrate the vast and inconsistent empirical findings.” (Kluger and DeNisi 1996)

Kluger and DeNisi propose such a theory, based on an extensive meta-analysis of prior research, termed ‘Feedback Intervention Theory’. The core premise is that, all else being equal, feedback interventions affect learner performance by directing the locus of their attention to one of three levels of control; a) task learning, b) task motivation and c) meta-task processes.

Figure 8: Kluger and DeNisi (1996) - Three levels of control



This theory proposes that, the lower down this hierarchy learners’ attention is directed, the higher effect of a feedback intervention on performance. Feedback intervention theory offers an alternative explanation to the issue of inconsistency in performance related research but its utility is compromised in that it reaches few generalizable conclusions and few predictions that are testable empirically. A fact acknowledged by Kluger and DeNisi.

“FIT in its present form lacks very detailed and specific predictions. In this form, FIT runs the risk of being unfalsifiable.” (Kluger and DeNisi 1996)

As yet, none of these explanations have been supported empirically.

Some researchers have questioned the underlying efficacy of the performance focused approach and called for new approaches to investigating TBI feedback (Mory 2003: 771, Handley, Price, et al. 2011: 548). This is discussed in the following section.

2.3 THE RESEARCH GAP

The previous section pointed to methodological issues which may explain the inconsistent results found in prior research relating to different attributes in the taxonomy of TBI feedback. One approach to progressing research in

this area would be to design a programme of studies aiming to address these issues. However, some have doubted the underlying efficacy of the performance focused approach and have called for research which places greater focus on the learner (Mory 2003: 771, Handley, Price, et al. 2011: 548).

Handley, Price, et al. (2011) point to two issues endemic to performance based studies. First, they propose the reductivist nature of prior research designs serves to ‘disaggregate’ the impact of each feedback attribute. In other words, that the impact of the interaction between the different attributes is not considered in such designs. Second, Handley, Price, et al. propose that such studies fail to consider the impact of learners’ engagement with feedback, suggesting this to be a confounding factor.

“(A) problem is that the atomistic examination of which variables cause which changes in test scores says nothing of the different ways that students interact with the feedback.”

(Handley, Price, et al. 2011: 546)

If the nature of learners’ engagement with TBI feedback impacts on the effectiveness of TBI feedback then this could explain the inconsistencies found in prior research. However, there have been very few studies focusing on learners’ engagement with feedback and this area remains largely unexplored.

Focusing on the learner is consistent with the constructivist paradigm. Since Mory’s (2003: 770) observation that there had been little research into TBI feedback from this paradigm, some research has focused on the learners’ perspective of educational feedback (Weaver 2006; Handley and Cox 2007; King, Schrodtt et al. 2009; Walker 2009). However, these studies mainly relate to the educational feedback received by learners from their teachers and not to TBI feedback. Weaver (2006) and Walker (2009) investigated learners’ perceptions of educational feedback received from teachers in the form of written comments on assignments whilst King, Schrodtt et al. (2009) researched learners’ perceptions of educational feedback received from teachers during classroom based instruction. Just one study was found to focus on TBI feedback, that of Handley and Cox (2007), which investigated learners’ perceptions of model-answers in a management development e-learning application.

Handley, Price, et al. (2011) argue for a research agenda focusing on learners’ engagement with educational feedback and highlight the absence of such research.

“Whilst there have been studies of students’ perception about the purpose of feedback (e.g. Walker, 2009), the process of student engagement with feedback has been relatively neglected. This is a serious omission if we accept the premise that the strongest influence on students is not the ‘teaching’ they receive but the assessment and feedback processes with which they engage.”

(Handley, Price et al. 2011: 548)

The present research contributes to addressing this gap by investigating learners’ preferences for the format of TBI feedback provided during technology based instruction. This research will ask, what format of TBI feedback will learners elect to view if given the choice? That is, which attributes across the taxonomy of TBI feedback will they elect to view if given the choice?

2.3.1.1 Note on research into learner control in computer based instruction

This literature review did not identify any prior research focusing on learners' preferences for the format of TBI feedback. However research into the notion of 'learner control' asks a related question, what are the results of giving learners control over the delivery of an activity? Research in this area is now discussed.

The nature of the control learners are provided within such research is broad. For example, Beck (2007) investigated the effect on retention of allowing learners to choose their own story to study rather than having one chosen for them. Ross, Morrison et al. (1989) investigated the effect of allowing learners to choose the context for practice examples in a course on statistics.

The particular aspects that a learner can control in such studies vary significantly but can be classified into four broad categories; a) control over the pace of an activity, b) control over the sequencing of individual components c) control over the content and d) control over the advice and support received. Studies in this area investigate the effect of control on learning, learners' attitudes and the length of time learners spend on a task (Kraiger and Jerden 2007: 67).

Studies investigating the effect of learner control on learning have been inconclusive. Niemiec, Sikorski et al. (1996) conducted a meta-analysis of 24 literature reviews which reviewed the effects of learner control on performance. They found a wide variation of effect size but when these were considered as a whole the mean effect was near zero (-.03). In a subsequent meta-analysis focusing on adult learners and considering changes in technology since Niemiec, Sikorski et al., Kraiger and Jerden (2007) found that learner control has a positive effect on learning although the mean effect size is small (0.11). This result likely reflects the significant variation in the nature of learner control in the studies that comprise the original data to these reviews.

This literature review identified three studies focusing specifically on learner control over feedback (Hansen 1974; Schimmel 1988; Pridemore and Klein 1991).

Hansen (1974) compared the effect of feedback on anxiety and performance across three groups of subjects. The first group received no feedback, the second group received KR plus IB type feedback and the third group were given the option of receiving KR and IB type feedback. Hansen found that learner control offers definite advantages both in terms of reducing learners' anxiety and increasing their performance. Similar reductions in anxiety were observed for the group who received KR plus IB type feedback and the group who had the option of doing so, however the latter achieved this result while requiring the subjects to do less work. Hansen, proposes that the additional information provided to the former group resulted in an increase in the 'error to problem' ratio.

It was not possible to obtain the report of the study by Schimmel (1988) during this literature review. The study by Pridemore and Klein (1991) compared four groups of undergraduate students studying psychology. All subjects studied the same material however the four groups differed in the level of learner control and the feedback provided. There were two levels of learner control. Under the first, subjects were automatically presented with feedback and under the second, subjects chose whether to receive feedback. There were also two levels of

feedback, first KR type feedback and second KCR plus EX type feedback. This study found that subjects who received KCR plus EX type feedback performed significantly better in a subsequent test than those who received KR type feedback regardless of the level of control. This study found no effect for learner control under either feedback condition.

2.4 FORMING HYPOTHESES

What hypotheses can be drawn regarding learner preferences relating to the attributes in the taxonomy of TBI feedback? What hypotheses can be drawn regarding the factors influence these preferences? These two questions are now addressed. As discussed throughout this section, prior research into the format of instructional feedback has focused on the correlation between feedback and performance in subsequent tests and it is primarily from this body of research that hypotheses may be drawn.

2.4.1 Which TBI feedback attributes in the taxonomy will learners chose to view given the choice?

This review of prior research resulted in the formation of three hypotheses relating to the frequency at which learners will choose to view the attributes in the taxonomy of TBI feedback. These hypotheses are now introduced and discussed.

2.4.1.1 Hypothesis one - Learners will elect to view feedback more frequently than they do not

Pridemore and Klein (1995) suggests that students are dissatisfied if feedback is not provided. As described above in the section ‘The content of feedback’, this study investigated the effects of TBI feedback on subjects’ performance as they studied an item of technology based instruction covering the use of a microscope. This study also investigated the effect of TBI feedback format on attitude which was measured using a Likert type survey covering topics such as enjoyment, motivation and confidence. The results indicated that the provision of TBI feedback had a significant positive effect on learners’ attitude.

“The level of feedback had a significant effect on student attitude toward the lesson. While most subjects were somewhat neutral about the lesson, subjects who did not receive feedback were significantly more likely to agree with the statements “I would have liked more information about my answers”, “I would have liked more information about the correct answers”, and “I would rather have the teacher give the lesson”. (Pridemore and Klein 1995: 449)

From this, it is hypothesised here that learners will elect to view feedback more frequently than they do not.

2.4.1.2 Hypothesis two - Learners will elect to view KCR and KR type feedback more frequently than they elect to view either EX, IB or EI type feedback.

On the basis of a meta-review of prior research by Bangert-Drowns et al. (1991) it can be deduced that users will elect to view KR type feedback. This meta-analysis reviewed 58 effect sizes from 40 reports and found that,

“There is ample evidence that, in instructional situations like the ones in these studies, learners use feedback to check the correctness of their answers.” (Bangert-Drowns, Kulik et al. 1991: 233)

The previously mentioned studies by Kulhavy, White et al. (1985), Phye (1979) and Dempsey and Litchfield (1993) suggest that learners will also elect to view KCR type feedback and also that they will view this feedback type more frequently than they elect to view either EX, IB or EI type feedback. These studies suggest the existence of a threshold beyond which the study of feedback becomes either ineffective or inefficient. They suggest that KCR type elaborative feedback represents this threshold and that attention to the other types of elaborative feedback are not worthwhile to the learner.

Pridemore and Klein (1995) offer further support for the existence of a threshold. In this study, subjects were presented with different feedback conditions including ‘no feedback’, KCR and KCR plus EX type elaborative feedback. The study found that subjects who received KCR plus EX type feedback were less likely to have enjoyed the lesson and less likely to feel they had ‘learned a lot’ than those receiving KCR type feedback alone.

From the above, it is hypothesised here that learners will elect to view KR and KCR type feedback more frequently than they elect to view either EX, IB or EI type feedback.

2.4.1.3 Hypothesis three - Learners will elect to view immediate feedback more frequently than they elect to view delayed feedback

There is an on-going debate as to whether feedback should be provided immediately following a learner’s response to a question or following a delay. Supporters of the delay-retention effect propose that delaying feedback will lead to increased retention of the subject matter. However, a number of researchers reject the general existence of such an effect (Kulik and Kulik 1988; Dempsey, Driscoll et al. 1993; Dihoff, Brosvic et al. 2003). This debate was discussed in detail earlier in this review.

For the purposes of investigation, it is hypothesised here that subjects will elect to view immediate rather than delayed feedback.

2.4.2 What factors potentially affect the TBI feedback types learners prefer to receive?

This literature review suggests nine factors affect learners’ selection of TBI feedback types. These nine factors can be organised into those that are internal to the learner and those which are external. Internal factors include gender

and cognitive factors such as response certitude. External factors are environmental factors relating to the design and delivery of the activity.

The nine factors are now presented and discussed starting with the internal factors which are given the prefix “#I”.

2.4.2.1 Factor #I1 Learner expectations

Handley (2003) found the key expectation held by learners regarding feedback to be:

“Learners expected feedback to enable them to achieve a sense of closure to the task.”

(Handley 2003: 206)

Thus, learners’ expectations of feedback potentially affect learners’ engagement with TBI feedback and the frequency at which they will elect to view a particular feedback type. Handley found that these expectations arise from four factors:

1. Learners’ ‘style of engagement’ with the instructional activity

Handley defines seven distinct styles ranging from ‘mindful learning’ to ‘going through the motions’.

2. Learners’ ‘sense of inquiry’ arising from the task

This refers to the learners’ level of interest in the activity.

3. Learners’ ‘assessment of personal risk’ arising from the task

This refers to learners’ assessment of their ability to deal with the real life scenarios presented in Handley’s study together with the implications of failure.

4. Learners’ generic expectations of feedback

Broadly speaking, learners generally expect feedback to verify, correct, and explain.

Learners’ ‘sense of inquiry’ also appears in the study by Dempsey and Litchfield (1993) described above in the section ‘Prior research investigating the content of TBI feedback’ (See page 50) . This study found that:

“A second question asked learners to choose which statement concerning immediate feedback was most true for them. The largest percentage of learners (46%) responded that it satisfied their curiosity about whether the answer was correct.” (Dempsey and Litchfield 1993: 11)

2.4.2.2 Factor #I2 Response certitude

As mentioned previously (See page 33), Kulhavy and Stock (1989) found ‘response certitude’, the degree to which learners are certain that a response is correct, to affect how long students spend on instructional feedback following an incorrect response. For example, if learners are certain they have answered a question correctly but are told they are incorrect then they will spend more time on feedback than if they had been less certain. Kulhavy and Stock

argue that response certitude strongly influences learners' expectations of instructional feedback and subsequent engagement with it.

2.4.2.3 Factor #I3 Learners' interpretation of the task

A number of models of feedback incorporate learners' interpretation of the task (Bangert-Drowns, Kulik et al. 1991; Kluger and DeNisi 1996; Handley 2003; Hattie and Timperley 2007). In their model of self-regulated learning, Butler and Winne (1995) suggest that learners' interpretation of a task is a factor directly affecting how learners use feedback. In this model, when faced with a task, self-regulating learners first use their knowledge and beliefs to construct an interpretation of what is required by a task. Based on this interpretation learners then set goals and determine the tactics and strategies required to achieve these goals. Upon receiving feedback self-regulating learners adjust these goals, tactics and strategies as required.

2.4.2.4 Factor #I4 Level of expertise

A number of studies indicate that learners' ability in a subject may impact on the effectiveness of feedback (Roper 1977; Butler and Winne 1995; Handley 2002).

In their model of self-regulated learning, Butler and Winne (1995) propose that a learner's level of expertise determines how that learner will use feedback. Butler and Winne view feedback as a core component of self-regulated learning. It is feedback, both that generated internally and that provided by external sources, that enables learners to evaluate and regulate their engagement with learning activities. Butler and Winne propose two reasons why learners' level affects their interactions with feedback. First, as learners' knowledge of a subject domain increases in depth and richness then their use of the cognitive strategies required for self-regulated learning also increases. One of these core cognitive strategies is the ability to use feedback to reinterpret elements of a task prior to engaging with it. Second, Butler and Winne state that higher level learners have a broader and deeper knowledge of the tactics and strategies required to complete tasks in the subject domain. Such learners are thus better able to adjust their approach if required following receipt of feedback. Butler and Winne (1995) suggest that learner level will impact on whether learners choose to view the different feedback types.

Hanna (1976) adds depth to this hypothesis by suggesting that low ability students are more likely to view elaborative feedback. This study investigated the effect of different types of feedback on learners' performance in a subsequent test. This study found that while verification type feedback was the most effective feedback type in increasing performance among high-ability students, elaborative feedback was most effective for low-ability students.

A further study, Roper (1977), suggests that learners' ability in a subject may impact on whether they elect to view immediate or delayed feedback. Roper found that delayed feedback to be more effective for learners with prior knowledge of a subject than for those studying for the first time. Thus, we can hypothesise that that lower level learners will be more likely to select immediate feedback.

One final study, Handley (2002), established that learners' ability in a subject affects the perceived usefulness of feedback. Handley interviewed 72 learners on their perceptions of feedback in an item of technology based instruction and found that learners' perception of the usefulness of different feedback types differed according to whether they were experts or novices in the subject.

From the above, it is hypothesised here that the extent to which users elect to choose the different attributes in the taxonomy of TBI feedback depends upon their level of expertise.

2.4.2.5 Factor #I5 Performance: Whether the learner answers a question correctly or incorrectly

As discussed in the section 'Feedback as viewed through educational theory' (See page 29), whether the role of feedback is to reinforce correct responses or to correct errors in incorrect responses was a contentious issue among educational researchers during the mid 20th century. However, research during this period failed to provide evidence that "positive feedback following positive responses acts in a reinforcing manner" (Mory 2003: 746). The results of a more recent study by Pashler, Cepeda et al. (2005), focusing on subjects' engagement with an item of technology based instruction, reaffirm this statement. Pashler, Cepeda et al. found that,

"Telling subjects they were right after they made a correct response produced no detectable improvement in either immediate performance or learning as assessed in a delayed test."

(Pashler, Cepeda et al. 2005: 7)

What can be deduced regarding learner preferences for TBI feedback from these results? It is hypothesised here that learners will not see the utility in electing to view TBI feedback following correct responses.

2.4.2.6 Factor #I6 Gender: The gender of the learner

Hanna (1976) found the effects of TBI feedback, provided following responses to multiple choice question based instruction, on subsequent post tests was more significant for males than females. From this it can be hypothesised that learners' gender affects whether they choose to view a feedback type.

2.4.2.7 Factor #I7 Error type: The type of error made by learners in answering a question

When learners respond to a question incorrectly, it can be hypothesised that the nature of the attention they give feedback depends on the type of error they made when selecting their response. This review identified three studies have investigated such a relationship (Tatsuoka 1984; Birenbaum and Tatsuoka 1987; Dempsey and Litchfield 1993).

Dempsey and Litchfield (1993), described above in the section 'Prior research investigating the content of TBI feedback' (See page 50), found 'error type' to be a determinant of the time learners spend on TBI feedback. This study categorized TBI feedback into four types and then conducted an experiment to establish whether feedback type was a determinant of the following: performance during instruction, performance in a subsequent test, time

spent on feedback and feedback efficiency. In addition, Dempsey and Litchfield used an error classification method developed by Driscoll and Tesser (1985 cited by Dempsey and Litchfield 1993) to classify the errors students as being either “fine discrimination errors” or “gross discrimination errors” and included this as a factor in their analysis. This categorization used frameworks from ‘concept learning’ (Bruner, Goodnow et al. 1967) which refers to a learner’s ability to organize items into different categories or concepts. For example the ability to recognize that a chair is a member of the category of furniture. Dempsey and Litchfield refer to three different types of error. First, ‘over generalization’ which occurs when a learner considers a non-instance to be an example of a concept. For example, considering a cup to be an item of furniture. Second, ‘under generalization’ which occurs when a learner considers an instance of a concept not to be so. For example, considering that a chair not to be an item of furniture. Third, ‘misconception’ which occurs when a learner makes both errors of overgeneralization and under generalization in their response. Dempsey and Litchfield found that,

“The combined feedback study times for fine discrimination errors were, on the average, almost twice as long as those for gross discrimination.” (Dempsey and Litchfield 1993: 11)

The findings of Dempsey and Litchfield are corroborated by Tatsuoka (1984) and Birenbaum and Tatsuoka (1987). These two studies also investigated the relationship between error type and the effect of TBI feedback on performance. Birenbaum and Tatsuoka (1987) investigated the interaction between error type, TBI feedback and performance in a subsequent post test. Birenbaum and Tatsuoka adopted three types of feedback in their study. KR, KCR and IB type feedback comprising the right rule for solving the problem. The study hypothesized that the feedback would remove ‘non-serious’ learner errors from the post test and also that none of the feedback types would remove ‘serious’ errors. The latter hypothesis was supported. None of the feedback types were able to remove serious errors. However, the study found that the effectiveness of the three feedback types in correcting non-serious errors varied significantly. The IB type feedback being more effective than simple KR type verification feedback. To explain this phenomenon, Birenbaum and Tatsuoka refer to the following statement by Kulhavy (1977),

“The feedback acts to inform the student concerning the accuracy of his response relative to knowledge he already possesses about the content. Obviously, if the material studied is unfamiliar or abstruse, seeing feedback should have little effect on criterion performance, since there is no way to relate the new information to what is already known.” (Kulhavy 1977: 220)

For the present study, while neither of these studies gives an indication of learner preferences for TBI feedback per se, they do suggest that error type may affect how learners choose to interact with TBI feedback.

That concludes consideration of internal factors that may affect the frequency at which learners elect to view the attributes in the taxonomy of TBI feedback. Two potential external factors are now introduced, these are assigned the prefix ‘#E’.

2.4.2.8 *Factor #E1 Assessment: Whether an activity is formally assessed*

A study by Morrison, Ross et al. (1995) found that learners' behaviour to be dependent on whether an activity is assessed. This study investigated the effects of assessment and different feedback strategies on performance. Subjects were 246 undergraduates enrolled on one of two compulsory introductory courses in teacher education. The testing instrument was an item of technology based instruction consisting of presentational material followed by multiple choice question based practice tests.

The subjects were divided into two groups. The first group was told that their performance in the technology based instruction would constitute part of their assessment for the course. The second group was told it would not. Each group was then sub-divided into five sub-groups, each of which was presented with a different format of feedback during the practice tests. The five feedback formats were 'no questions presented', 'no feedback presented', 'delayed knowledge of correct response', 'knowledge of correct response' and 'answer until correct'. Both groups were given the opportunity to view 'review screens' during feedback.

Morrison, Ross et al. found that subjects for whom the instruction constituted part of assessment performed significantly better than the second group and did so across all five feedback types. In addition, the assessed group was 3.24 times more likely to view the review screens following an incorrect response.

Given this result, it is hypothesised here that learners' preferences for TBI feedback types depend on whether the activity being performed constitutes part of the course assessment.

2.4.2.9 *Factor #E2 The level at which feedback is directed*

Kluger and DeNisi (1996) and Hattie and Timperley (2007) both make reference to the fact that feedback can be directed at different levels. While the focus of these models differs, both agree that feedback focused at the task level, that is focused specifically on the outcome or product of the task, is more effective than feedback focused at higher levels such as general praise for a learner.

From this, it can be hypothesised that learners are more likely to choose feedback that relates specifically to a task.

2.5 LITERATURE REVIEW: SUMMARY

This literature review first outlined how the role of feedback is considered within the three dominant paradigms in educational research and then discussed five models of instructional feedback that potentially act as theoretical frameworks through which to interpret the results of this study.

The taxonomy of TBI feedback attributes adopted in this research was then produced following an analysis of prior classifications of TBI feedback. This taxonomy comprises types, load and timing attributes and draws a distinction between verificational feedback and elaborative feedback. Verificational feedback informs the learner as to the

correctness of their response. Elaborative feedback is additional information that may guide the learner towards specified learning objectives. The 'format' of TBI feedback was then defined to be a particular selection of the attributes in this taxonomy.

Prior research investigating the effect of differing formats of TBI feedback on learners' subsequent performance was then reviewed and it was found that the relative effectiveness of different formats of TBI feedback remains unclear despite a large number of studies in this area. A number of methodological issues offering explanations for this inconsistency were introduced but it was noted that none of these explanations have been or are able to be tested empirically.

The review then introduced studies that question the underlying efficacy of the performance focused approach and call for research which places greater focus on the learner (Mory 2003: 771, Handley, Price, et al. 2011: 548). An improved understanding of learners' engagement with TBI feedback could allow consideration of engagement to be incorporated in studies investigating the effectiveness of TBI feedback and therefore has value as a focus of research. The review noted a paucity of such research and this was identified as the broad gap in the literature to which the present research can contribute. Specifically, this research will focus on learners' preferences for the format of TBI feedback. These preferences will be measured by the frequency at which learners choose to view different attributes of TBI feedback when given the choice.

The review concluded by analysing the results of prior research to form hypotheses for investigation. Three general hypotheses regarding the frequency at which learners will elect to view the different feedback attributes were formed.

1. Learners will elect to view feedback more frequently than they do not
2. Learners will elect to view KCR and KR type feedback more frequently than they elect to view either EX, IB or EI type feedback.
3. Learners will elect to view immediate feedback more frequently than they elect to view delayed feedback

In addition, nine factors were identified that potentially affect whether learners choose to view a TBI feedback type. These were categorised into factors that are internal to the learner and those which are external. Internal factors include gender and cognitive factors such as response certitude. External factors are environmental factors relating to the design and delivery of the activity.

3. RESEARCH METHODOLOGY

3.1 THEORETICAL FRAMEWORK

This section outlines the theoretical frameworks adopted in this research. The overarching framework through which the results of this research are interpreted is constructivism. Constructivist thought informs assumptions regarding the nature of reality and knowledge, the nature of the learner and the nature of learning. Two additional frameworks from the field of cognitive psychology are adopted to add detail to the latter, schema theory and cognitive load theory.

These frameworks are now discussed in turn.

3.1.1 Constructivism

3.1.1.1 Clarification of the term ‘constructivism’ as adopted in this research

As discussed in the section “Feedback as viewed through educational theory” in the literature review (See page 29), Duffy and Cunningham (1996: 181) draw a distinction between the ‘cognitive constructivist’ and ‘socio-constructivist’ uses of the term ‘constructivism’. The former focusing on the individual and the processes within the mind of a learner through which learning occurs. The latter emphasising the social nature of knowledge formation and the importance of the context in which learning occurs. The present research focuses on self-directed activities and the theoretical framework adopted draws specifically on concepts from ‘cognitive constructivist’ theory. It is thus ‘cognitive constructivism’ that forms the basis for the theoretical framework.

3.1.1.2 On the nature of reality and knowledge (Ontological and epistemological considerations)

This research aims to extrapolate knowledge relating to learners’ preferences for TBI feedback. In order to design a robust research approach the underlying ontology (relating to the nature of existence and reality) and epistemological (relating to the nature of knowledge) assumptions regarding the nature of this knowledge require clarification. First, the distinction between ontological and epistemological considerations is clarified.

Ontological considerations

Ontological considerations concern the nature of reality. A dichotomy runs throughout ontological thought. First, there is the objectivist or realist proposition that proposes the existence of an objective reality that exists outside of the mind. In this position, the external world around us represents the absolute truth and learners’ knowledge of the world is an approximation of this truth. Much educational research in the 20th century, for example that based on behaviourism and information processing theory, assumes such an objectivist position. In contrast, the

relativist or nominalist position proposes reality only exists in the mind. In this view, all aspects of reality are interpreted through human perception and consideration.

Epistemological considerations

Related to ontology is epistemology, which concerns the nature of knowledge. That is our knowledge of the reality considered by ontology. Consideration of epistemology is required in this research as the subject under investigation concerns one aspect of the process of learning. One fundamental component of learning is the acquisition of knowledge.

In research on matters of education, epistemology is often considered in terms of two paradigms, the positivist paradigm and the interpretivist paradigm. The positivist position asserts that authentic knowledge can only be based on sensory experience and advanced through observation and experimentation. The origins of the term can be back to the writings of the 19th century philosopher, Auguste Comte (Cohen, Manion et al. 2007: 9). Comte was working within a tradition of empiricism which placed itself in opposition to prevailing religious and philosophical beliefs that proposed a non-empirically based view of the world (Pring 2004: 91). In a significant move, the empiricists abandoned metaphysical attempts to create knowledge through reason alone and focused only on that which can be observed. Comte's contribution was to apply this positivist approach beyond the realm of the physical world and into the sphere of society and human affairs. This resulted in the beginnings of a 'science of society' and in the present day positivism is still strongly associated with the scientific method of enquiry.

Interpretivism is used to define an epistemological viewpoint in contrast to positivism. As such, the term is used broadly to include the views of those who feel that the subject matter of the social sciences, human beings and social structures, is fundamentally different to that of the natural sciences and as such cannot be the subject of scientific enquiry.

Ontological and epistemological considerations: The constructivist perspective

Constructivism is associated with a relativist rather than realist ontological position. In matters of epistemology, supporters of the constructivist paradigm articulate their epistemological standpoint in contrast to the positivist position. As such, constructivist views on epistemology are interpretivist in nature. Ernst Von Glasersfeld puts the constructivist view thus,

“As for Vico, knowledge for Piaget is never (and can never be) a 'representation' of the real world. Instead it is the collection of conceptual structures that turn out to be adapted or, as I would say, *viable* within the knowing subject's range of experience.” (Von Glasersfeld 1989: 125)

Proponents of the constructivist paradigm propose that learners' knowledge does not necessarily bear any relation to an external reality and does not need to in order to be useful and viable. From the constructivist standpoint, the truth of any knowledge is dependent upon a knower's frame of reference.

3.1.1.3 *The nature of learning*

Constructivism focuses on the processes learners undertake during learning. New knowledge is obtained when learners encounter an experience that challenges their existing representation of the world. This is termed a perturbation. Learning occurs as learners resolve this perturbation and existing representations are refined to accommodate the new experience. This process is described by Ernst Von Glasersfeld (1989) as follows,

“Knowledge is never acquired passively, because novelty cannot be handled except through assimilation to a cognitive structure the experiencing subject already has. Indeed, the subject does not perceive an experience as novel until it generates a perturbation relative to some expected result. Only at that point the experience may lead to an accommodation and thus to a novel conceptual structure that re-establishes a relative equilibrium.” (Von Glasersfeld 1989: 136)

Thus learning occurs as learners regain equilibrium following such a perturbation. Von Glasersfeld explains the nature of this equilibrium as follows,

“(In the sphere of cognition) equilibrium refers to a state in which an epistemic agent's cognitive structures have yielded and continue to yield expected results, without bringing to the surface conceptual conflicts or contradictions.” (Von Glasersfeld 1989: 126)

This concept of learning occurring through perturbations which challenge cognitive equilibrium leading to an accommodation and a return to equilibrium, is a key component of the theoretical framework adopted to interpret the results of this research. This concept has precedent in existing models of feedback. In Hattie and Timperley's (2007: 82) model, the role of feedback is to close any gap between what is currently understood by a learner and what is aimed to be understood. This concept of a gap builds on the work of Sadler (1989) who posits three indispensable conditions for learning. Sadler states that, in order to improve performance, a learner needs to

- a) Possess a concept of the standard (or goal, or reference level) being aimed for
- b) Compare the actual (or current) level of performance with the standard, and
- c) Engage in appropriate action which leads to some closure of the gap.

(Sadler 1989: 121)

In this research, perturbations are said to cause 'knowledge gaps' which learners then seek to accommodate. The role of TBI feedback in promoting learning is to assist learners in identifying and resolving these knowledge gaps.

3.1.1.4 *The nature of the learner*

In this research, constructivism also provides the framework through which to view learner behaviour. Building upon the constructivist theory of knowledge, educational researchers have formed a constructivist view of the

learner and of learner behaviour. Constructivism sees learners, not as passive participants, but as the primary active agents in the learning process.

“Knowledge is constructed by learners as they attempt to make sense of their experiences. Learners are not empty vassals waiting to be filled but are active organisms seeking meaning”

(Driscoll 2000: 376).

A corollary of this view is a recognition that learners themselves ultimately determine what learning is acquired from an educational experience. Learners are considered to be rational, self-directed and able to learn independently.

“In the constructivist model ... it assumes that young people are rational decision-makers, can be self directed and learn autonomously. Learners are encouraged to make choices about their learning experiences, within limits placed on them by teachers and the school context.”

(Askew and Lodge 2000: 9)

Thus constructivist assumptions form the nature of reality and knowledge, the nature of learning and the nature of the learner that form the basis of the theoretical framework adopted in this research.

3.1.2 Schema theory

The term ‘schema’ emerged from the work of Sir Frederic Bartlett in the early 20th Century (Anderson and Pearson 1984: 256) and the application of the concept of schemas to educational research can be traced to the work of Anderson, Spiro et al. (1978). A schema refers to a representation of knowledge within human memory. Any collection of knowledge within memory is considered to be a collection of schemas. The process of learning involves incorporating new knowledge into existing schemas or, if no relevant schema exists, generating new ones. Schema theory is concerned with identifying the processes through which schemas are revised or generated.

A central tenet of schema theory are the notions of long-term memory and short-term memory, also called working memory (Derry 1996). The relationship of these two concepts to schemas is described by Derry as follows.

“(Schema theory) psychologists believe in long-term memories that store previously learned schemas, and working memories that represent a person’s span of immediate attention. Thinking and learning take place within working memory, where prior-knowledge schemas are activated in response to environmental input, providing context for interpreting experience and assimilating new knowledge.”

(Derry 1996: 167)

This notion of a requirement for schemas to be activated and placed into working memory is adopted in the interpretation of results from the present research.

3.1.3 Cognitive load theory

Cognitive load theory emerged from cognitive psychology and focuses on the role played by a limited working memory in processing information. Cognitive load theory builds on tenets of schema theory including the notions of long-term and short-term memory and that knowledge is organised within memory in the form of schemas (Sweller 1988: 257, 264).

Cognitive load theory posits that there is a limit to the volume of schemas that can be acted upon within working memory and investigates how this load can be managed in order to increase the efficiency of learning. As Chandler and Sweller (1991) state,

“Cognitive load theory is concerned with the manner in which cognitive resources are focused and used during learning and problem solving. Many learning and problem-solving procedures encouraged by instructional formats result in students engaging in cognitive activities far removed from the ostensible goals of the task. The cognitive load generated by these irrelevant activities can impede skill acquisition.” (Chandler and Sweller 1991: 294)

Sweller, van Merriënboer et al. (1998: 259) propose cognitive load can be categorized as being either intrinsic, extraneous or germane. Intrinsic load is determined by the inherent level of difficulty in the material being studied and reflects that fact that some concepts are more difficult to acquire than others. Extraneous cognitive load is determined by the manner in which materials are presented to the learner. Germane cognitive load refers to the work done by working memory in the construction of schemas. Sweller, van Merriënboer et al. propose that changing the format of instruction can lead to more effective and efficient learning experiences by decreasing extraneous cognitive load and increasing germane cognitive load. Investigating the mechanisms by which this can be achieved forms the focus of research into cognitive load theory.

3.1.4 On the cohesion of the theoretical framework

Given that schema theory and cognitive load theory emerged from a cognitivist perspective of learning, the question arises as to whether these theories are compatible with the constructivist view that forms the overarching framework for the present research. This research contends there is no contradiction between these theories as they operate at different hierarchical levels of analysis. Rather, they support each other with schema theory and cognitive load theory adding greater detail to the notion of learning proposed within constructivism.

With regard to schema theory, this view is supported by Derry (1996) who sees no conflict between this theory and the constructivist view and notes how the notion of schemas has been “assimilated into the rhetoric” (Ibid 1996: 167) of constructivist theory. Derry argues that schema theory is not attached exclusively to either the cognitivist or constructivist views of learning but is compatible with both and has the potential to act as a common ground for exponents of each orientation (Ibid: 172).

A similar argument supports the inclusion of cognitive load theory. As described above, much of the focus of cognitive load theory has focused on the delivery of instruction to learners and how this can be designed to promote efficient learning. This focus on what educators do seems incompatible with the constructivist focus on the learner however, as Valcke (2002) notes,

“This critique does not touch the credibility of (cognitive load theory) to explain and describe cognitive processes.” (Valcke 2002: 2)

It is therefore proposed that the adoption of schema theory and cognitive load theory within an overarching constructivist paradigm forms a cohesive theoretical framework and one that has precedent within educational research (Mayer, Moreno et al. 1999; Mayer and Moreno 2002).

3.2 RESEARCH DESIGN

A research design can be defined as,

"... a framework for the collection and analysis of data." (Bryman and Bell 2003: 32).

This section describes how an exploratory design consisting of computer based observations and semi-structured interviews was chosen to investigate the gap identified in the literature review. This section describes why these methods were adopted and discusses alternative strategies that were considered. A following section 'Research methods' provides detail on the methods adopted to collect and analyse data in this research.

This section first presents a discussion of the two dominant strategies within social science research and then reviews a number of potentially relevant research designs considering the relevance of each to this research. Following this, the final research design is presented.

3.2.1 Research strategy: Quantitative vs. qualitative research

A discussion of research strategy within the social sciences, including educational research, needs to address the perceived duality in which designs are classified as being either 'quantitative' or 'qualitative'. At a surface level the difference between these two strategies is straightforward in that, under the former, researchers identify quantifiable constructs relating to the object under investigation, measure these constructs within a particular context and then perform an analysis of the resulting measurements. In qualitative analysis, such numerical measurements are rarely performed. However a consideration of the underlying ontological (relating to the nature of existence and reality) and epistemological (relating to the nature of knowledge) assumptions behind the two strategies has led to vigorous debate between proponents of opposing strategies (Pring 2004).

3.2.1.1 *The quantitative paradigm*

The quantitative research paradigm can be broadly defined as:

"Entailing the collection of numerical data and as exhibiting a view of the relationship between theory and research as deductive, a predilection for a natural science approach (and of positivism in particular) and as having an objectivist conception of social reality." (Bryman and Bell 2003: 68)

As Bryman and Bell state, the quantitative paradigm is usually closely associated with an underlying positivist epistemology. As described in the previous section 'Theoretical framework', the positivist position asserts that authentic knowledge can only be based on sensory experience and advanced through observation and experimentation. The quantitative method determines numerical constructs relating to human and societal characteristics and then seeks to determine the relationship behind such constructs or identify factors that explain observed changes in their value. A key feature of much quantitative research is that it seeks to identify findings that

are generalizable across a larger societal group. It is this generalizability that makes quantitative research particularly valuable to fellow researchers and attractive to research funders and government policy makers.

However, the quantitative approach has come under much criticism (Pring 2004: 50-51; Cohen, Manion et al. 2007: 17). Much of this criticism is due to one of two underlying issues. First, opponents of the positivist quantitative paradigm assert that we human beings cannot be the objects of science due to our idiosyncratic nature. Second, that reality relating to humans and social groups rests within the mind and is unique to each individual. Opponents of the positivist quantitative paradigm point to risk of assuming an epistemological standpoint based on a 'naïve reality' which fails to consider that human individuals' perception of reality is subjective.

In response, proponents of the positivist quantitative paradigm assert that there should be no clear distinction between research into the physical world and that into the social world. They propose that there are social facts that can be characterized, quantified and analysed as per their counterparts in the physical world. This view supposes that, despite the differences that exist between individuals, commonalities do exist and these can be identified and measured. Likewise, the view holds that the infrastructures, systems and cultural behaviours that comprise society also constitute identifiable facts and are therefore a legitimate focus for scientific research. Pring (2004: 52) argues that in the process of reality being negotiated between individuals it takes on characteristics that may be subject to quantitative research.

3.2.1.2 The qualitative paradigm

The qualitative paradigm evolved to address the criticisms of the quantitative paradigm discussed above and as such can be described in contrast to it. First, the data obtained tends to comprise words as opposed to numbers. Second, the generation of theory tends to be inductive rather than deductive. Third, the epistemological position tends to be interpretivist rather than positivist. Finally, the ontological position of the qualitative paradigm is relativist rather than objectivist meaning that an emphasis is placed on reality as constructed within the minds of research subjects as opposed to an objective 'real world' external to the mind (Bryman and Bell 2003: 280).

Qualitative researchers emphasize the subjective nature of events in the mind of individual research subjects and aim to view such events through the eyes of the people under study. In doing this they are able to elicit information relating to the subjects' perception of an event and the intention behind any observed behaviour.

A criticism of qualitative research is the degree to which results are generalizable (Cohen, Manion et al. 2007: 25). If reality is subjective and unique to an individual, then research findings cannot be generalizable and are as such limited in value.

3.2.2 The fallacy of the duality and the need for a mixed approach

A pragmatic view is that, while rationales exist for each paradigm, both can be considered valid. As such the choice before the researcher is not 'either or' but to select methods from each paradigm according to their suitability to the research task at hand.

Justification for a pragmatic approach is that too sharp a distinction can be drawn between the two paradigms and that this distinction does exist in practice. For example, the argument proposed by proponents of the qualitative paradigm supposes that researchers adopting quantitative methods are either not aware that knowledge can be constructed by subjects or are aware and fail to take this into account. Pragmatists propose that the solution to potential problems of this nature is to adapt rather than abandon the quantitative approach.

Pring (2004) advocates a pragmatic approach and, in the following extract, suggests that a duality does not need to exist in practice:

“Attention to the many ways in which research is conducted reveals a more complicated picture. In some respects, people are the object of science – of generalizations and causal explanations. In other respects, however, they escape such explanations through interpreting the world in their own personal ways. Different methods get at these different explanations. Understanding human beings, and thus researching into what they do and how they behave, calls upon many different methods, each making complex assumptions about what it means to explain the behaviours and personal and social activities.” (Pring 2004: 56-57)

This research adopts this position proposed by Pring and assumes that both the reality described by the positivist paradigm and that described by the interpretivist paradigm co-exist. In other words, these two realities are not independent of each other but in fact closely related. A pragmatic view is taken regarding the dichotomy between qualitative and quantitative methods.

3.2.3 Designs considered

This section reviews a number of potentially relevant research designs and discusses their potential value to the present research

A notable characteristic of research relating to matters of education is the variety of research designs that have been adopted and Pring (2004: 33) proposes that this is due the complexity of the subject matter under investigation, namely the process of learning. This complexity results in a wide variety of problems being researched and as a result a wide variety of designs need to be available to the researcher.

The following list of research designs of potential relevance to the present research derives from consideration of designs appearing in the literature review (See page 46) and an analysis of three textbooks on research in the social sciences (Bryman and Bell 2003; Pring 2004; Cohen, Manion et al. 2007) . The appropriacy of each design to addressing the research questions was considered and this is reported below. A full consideration of the advantages and disadvantages of each design is beyond the scope of this report however a brief summary of each approach is given followed by a discussion of those issues of most relevance here.

3.2.3.1 Observation

In discussing the observational approach Pring (2004: 33) states simply that, if one wants to know something, then the a simple approach is to go out and take a look. An observational design can be summarized as,

“The systematic recording of observations and the attempt to generalize from them.” (Pring 2004: 33)

If the resultant generalizations are confirmed in repeated observations and also established by third party observers then the generalizations may be adopted as theory for use by practitioners and the research community. In this way, theory is developed inductively, through exploration.

In studies of human behaviour such as the present study, observation can be considered more reliable than alternative approaches such as surveys and interviews in that the subject is not asked to interpret their behaviour. Rather the researcher observes that behaviour directly.

However there are a number of disadvantages with this design. A typical observational design entails the researcher observing the actions of a subject, such as a teacher or student, as they perform a particular task. Such a design is clearly time intensive, particularly if a large number of subjects are observed. Having more than one observer is a possible solution to this however this then raises the issue of consistency across the observers. A further issue is that this method suffers from the problem of reality being filtered through the interpretation of an observer.

For the present study, these issues can be mitigated to a significant degree by developing computer software to track learner behaviour and act as an objective surrogate observer.

If we want to identify learners' preferences for TBI feedback then observing learners' interactions with such feedback is an obvious approach. However, observation alone may not answer the question to the extent required due the nature of the data that are obtained. Data obtained from observation alone are not likely to provide explanations for any findings, for example learners' intentions behind their actions. In addition, repeated observations will be required in order to discover whether any findings are consistent.

3.2.3.2 Experiments

Experiments are the most frequent design adopted in prior research into TBI feedback. It was initially imagined that an experiment would also form the basis for the design of this research. Two types of experiment were considered; true experiments and quasi experiments.

True experiments

Experiments are a type of observational research, albeit a very structured one. Experiments attempt to establish causality between one variable and another. Using observations and measurement, they attempt to prove that if a variable, X, is manipulated then a further variable, Y, changes as a result. In this case, X is termed the independent variable and Y is termed the dependent variable. In order to establish such causality, experiments pay a great deal

of attention to the conditions in which observations are made. In particular, researchers attempt to remove any factors aside from X that may also be the cause of any changes in Y.

In the classic true experimental design (Cohen, Manion et al. 2007: 275), two subject groups are first established, the control group and the experimental group. The experimental group receives the intervention that is the subject of investigation. The control group does not receive the intervention and acts as a baseline from which any changes in the experimental group can be assessed. Subjects are assigned randomly to one of the two groups. The dependent variable is measured before and after the intervention so that before-and-after analyses can be conducted. An experiment of this type is referred to as a randomized controlled trial or RCT.

A key advantage of experiments is that they tend to have high reliability and validity.

“Experimental research is often held up as a touchstone because it engenders considerable confidence in the robustness and trustworthiness of causal findings.” (Bryman and Bell 2003: 39)

While strong in terms of validity and reliability, true experimental designs suffer from a number of drawbacks in the social science setting. These are primarily related to the need to remove subjects from a natural context and into an artificial experimental one. First, it may not be practical or ethical to place subjects into an experimental environment. Second, subjects may change their natural behaviour as a result of being placed into the experimental environment, a phenomenon known as the Hawthorne Effect (Bryman and Bell 2003: 45). Finally, the structural requirements for pre-tests of the dependent variable followed by an intervention followed by post-tests may lead to unnatural ‘interaction effects’ between these elements. For example, exposure to the pre-test may affect subjects’ reaction to the intervention or behaviour during the post-test.

A limitation of the experimental method is that, while the method can establish the cause of effects, it does not tell us why or how effects occur. In other words, experiments are effective at ‘causal description’ but not at ‘causal explanation’ (Bryman and Bell 2003: 16). To address the latter, additional methods are required.

For the present research, a significant disadvantage of the true experimental design is that the structural requirements compromise the authenticity of the study environment. This in turn compromises efforts to ensure that subjects behave in a natural manner and therefore the external validity of the design.

Quasi Experiments

In a social setting the need to control the experimental environment and remove possible confounding factors presents obvious problems due to the complexity of the social interactions and individuals involved. As Pring (2004) states,

“(In the social context) there can never be the purity of the scientific world where standard and limited conditions can be assured.” (Pring 2004: 65)

As a result, true experiments can be problematic in areas of social science such as educational research. An alternative but closely related approach is that of quasi-experiments. Quasi experiments differ from true experiments in that one or more of the requirements of the classic experimental design described earlier are not present (Cohen, Manion et al. 2007: 282). Usually, this refers to an experimental design in which subjects are not randomly assigned to the control group and the experimental group. Such a compromise is common in educational research as this enables a researcher to study groups that are already intact, for example an existing cohort of students. Quasi experiments can be categorised into four different types; pre-experimental designs, natural experiments, field experiments and non-randomised quasi experiments.

In pre-experimental designs, the characteristics of a true experiment are present aside from the presence of a control group (Ibid.: 282). A single group, such as a class of students, is studied but no comparison is made to an equivalent group.

In natural experiments, researchers do not obtain data through directly manipulating and observing variables but attempt to obtain such data from naturally occurring events in the real world (Bryman and Bell 2003: 45).

In a field experiment, researchers control and isolate independent and dependent variables but do so in a natural setting, for example a classroom, rather than in a laboratory. For this research, a field experiment has the advantage that any findings will result from a realistic context. The key disadvantage of field experiments is that it is more difficult to control for confounding factors and thus the internal validity of any claims regarding causality can be compromised.

Non-randomised quasi experiments differ from true experiments in that subjects are not randomly assigned to the control group and the experimental group. For the present research such experiments have the key advantage that existing cohorts of learners can form the research groups.

A drawback of quasi experiments is that the experimental group and control group may have different characteristics prior to an experimental treatment thus creating the possibility that any observed effects be caused by extraneous factors. As a result, the researcher needs to be diligent in considering such alternative explanations in order to estimate the treatment effect.

“In quasi-experiments, the researcher has to enumerate alternative explanations one by one, decide which are plausible, and then use logic, design, and measurement to assess whether each one is operating in a way that might explain any observed effect.” (Andre and Thieman 1988)

The process of determining which alternative explanations to consider also presents challenges. It is generally not possible or practical to consider all logically possible explanations. Researchers focus on the subset of logically possible explanations which are also plausible. Whether an explanation is plausible depends on ‘social consensus, shared experience and empirical data’ (Andre and Thieman 1988). The focus on plausible explanations does leave resulting knowledge claims vulnerable to subsequent discovery of casual variables which had been considered implausible.

The main disadvantage of quasi experiments is that, in comparison with true experiments, they have a lower degree of internal validity. That is, we cannot be as certain that any findings suggesting causality between the independent and dependent variables are real. However, quasi experiments do have the advantage that they can be more practical. In some instances, true experiments may not be possible and quasi experiments may be the only option available. In addition, compromises regarding internal validity may be offset by increases in external validity. That is, the fact that quasi experiments often occur in natural environments can increase the degree to which findings can be generalised.

3.2.3.3 Surveys

In this design, research subjects are asked to complete a questionnaire or survey. Surveys offer a number of advantages over other research approaches and are widely used in the social sciences (Cohen, Manion et al. 2007: 205). One key advantage is that data regarding the actions or intentions of the research subjects are provided by the subjects themselves and not interpreted by an observer. Surveys can be less time intensive than alternatives such as observation or qualitative interviews and this allows a larger number of subjects to be researched. A further and significant advantage is that surveys can provide quantitative data that can be subject to statistical analysis. For the present research, surveys may be an effective method to obtain data relating to learners' intentions behind their actions.

However, surveys also suffer from a number of disadvantages. As discussed above, in observational methods a key issue is that data need to be interpreted through the mind of the observer. This is less an issue when surveys are used however a parallel problem exists in that questions are interpreted through the mind of the subject (Pring 2004: 38). Subjects bring their own distinctive way of looking at the world to bear when replying to questions in a survey.

For the present research, a further problem is that of recall. Subjects are unlikely to be able to recall the detail their TBI feedback preferences to the degree required to answer the research question.

3.2.3.4 Interviews

In research interviews, subjects typically answer a series of questions posed by a researcher that relate to the research question.

Structured interviews are a type of interview commonly used in social science research and subjects within this field such as educational research. A structured interview is one in which the interview stimuli, typically questions, are determined in advance and change very little during the interview (Bryman and Bell 2003: 116). Such interviews have the advantage that each respondent is presented with exactly the same stimuli and as such responses can be aggregated.

The designs discussed in this section so far have either relied on an observer to interpret the actions of research subjects or restricted subjects' own description of their actions to the confines of a structured questionnaire.

Interviews have the key advantage of allowing the subjects more freedom to explain their actions and to communicate their perspective. The process can reveal greater insights relating to subjects' actions and the intentions behind them.

That said, the interview approach also suffers from some disadvantages. One such disadvantage is a possible loss of objectivity in that data obtained are first a product of the subjective interpretation of the subject and then subsequently a product of further interpretation by the interviewer (Pring 2004: 40). A more practical disadvantage is that interviews are time intensive. This restricts the number of possible research subjects which in turn can have a negative impact on the generalizability of any findings.

For the present research, a key problem in adopting interviews as the primary research tool is that the data collected may be inaccurate due to the difficulty subjects may have in accurately recalling the detail of their actions. However while interviews may not be an effective method of recording learner preferences, they may be an effective tool for revealing insights that assist with the interpretation of data collected via a different approach. In particular insights that relate to subjects' reasoning behind their choices.

3.2.3.5 Case studies

A criticism of all the designs above is that they seek to obtain results generalizable across a group of subjects whereas human subjects and the context in which they are situated are too unique for this to be possible (Pring 2004: 40). A solution to this issue is the case study approach in which a researcher presents a comprehensive study of one subject detailing a personal perspective and providing information regarding the context from the subjects' perspective.

For the present research, the inability of the case study approach to produce findings that are generalizable is a key failing in that such findings will both fail to answer the research questions identified and also fail to achieve the stated aims regarding a contribution to theory and practice.

3.2.3.6 Historical research

Historical research can be defined as,

“The systematic and objective location, evaluation and synthesis of evidence in order to establish facts and draw conclusions about past events.” (Cohen, Manion et al. 2007 citing Borg 1963: 191)

Historical research is not simply secondary research, in which a researcher adopts existing data, but attempts to adopt a historical mindset when considering the data. This means that the researcher carefully considers the human mindsets and environmental context in existence when the data were generated.

An historical approach to the present research would provide a longitudinal basis to findings with potentially strong reliability and validity. However, successfully obtaining access to relevant historical data does not appear possible

for two reasons. First, the literature review reveals an absence of published research into learner preferences. Second, no possible primary sources of such data are known to the researcher.

3.2.3.7 Ethnographic research

The ethnographic approach is a type of observational research with origins in anthropology (Bryman and Bell 2003: 315). In anthropology, ethnography is used to learn more about the society and behaviour of undocumented cultural groups. A researcher typically spends months or years living with the target group, observing behaviour and taking notes before returning home to document what has been discovered. For example, ethnographic techniques are used within management studies to identify the cultural aspects associated with a particular firm. Ethnography is used in educational research to identify cultural factors relating to particular student or teacher groups.

The focus of the present research is primarily on learners' interactions with a computer based instruction. Ethnographic techniques are primarily concerned with culture. While culture may play a role in influencing learners' preferences regarding feedback, and will be considered, this will not be the primary focus of the research. As such, ethnographic techniques are unlikely to be an efficient method of obtaining relevant data and will not be adopted.

3.2.3.8 Phenomenography

A criticism of observational types of research and restricted response types of research such as surveys is that they fail to address the complexity of interactions within a social context and also the unique worldview and subjective experience of an individual human subject (Pring 2004: 98). Phenomenographical techniques seek to address this issue. Phenomenography is a type of interview based research which focuses on subjects' descriptions of the event under investigation.

The focus of phenomenography is on the subjective experience of the subjects under research (Pring 2004: 100). In the educational setting, phenomenography has been used to gather information on how particular students experience an educational environment. For example, the experience of female students entering a primarily male educational environment.

For the present research, knowledge of how learners experience feedback will prove useful during the process of interpreting quantitative data relating to their preferences. However, a closer focus on the specifics of action and intention relating to feedback choices is of more direct relevance to the research questions. Given this, structured interview techniques are likely to produce more directly relevant data than phenomenographical techniques.

3.2.3.9 Action research

In the educational setting, action research involves teaching practitioners actively reflecting on their teaching with a view to its improvement (Cohen, Manion et al. 2007: 297). Action research differs from traditional academic research in that its aim is not to produce new knowledge but to improve practice.

Action research can be considered a form of critical theory in that it seeks to critique and change aspects of society rather than explain and understand them. From this standpoint, the approach is incompatible with the primary aims of the present research which are to understand and explain aspects of learner behaviour. In addition, by its nature, action research tends to focus on the teacher rather than the learner. For example, investigating how changes in teaching practice can improve educational outcomes. This too is incompatible with the present research which focuses, not on the teacher, but on the learner.

3.2.4 The final research design

The review of research designs, outlined above, argued the use of both observation and semi-structured interviews to be most appropriate to this research and most able to contribute towards answering the research questions. This section describes how a mixed research design incorporating both methods was adopted as the final research design. First, the rationale for not adopting an experimental design is stated.

3.2.4.1 The rationale for not adopting an experimental design

During the initial consideration of the research design, it was imagined that an experiment of some kind would form the basis of the design. Such an approach would be in keeping with prior research in this field and the experimental approach offered the advantage of high validity.

After further investigation it was concluded that the issue of validity was more complex. A true experiment would have high internal validity however the requirement to remove students from the authentic learning context would compromise the external validity of any results. The issue when conducting quasi experiments is reversed. Such experiments potentially offer high external validity but great care is required to ensure an acceptable level of internal validity. Such care includes measures to ensure that the control group is similar enough to the experimental group to act as an effective control. In addition, possible confounding variables need to be identified and controlled for, usually through the additional testing or questioning the members of both groups. The early decision to research existing students studying authentic learning tasks resulted in this research taking place within the restrictions of an authentic study environment. As a result it was not possible to make the interventions necessary to ensure the internal validity of any quasi experiment was acceptable.

However, the issue of validity was not the only determining factor in the decision not to proceed with an experimental approach. Experiments fall within the category of deductive research. Researchers first form a hypothesis, for example by forming an educated guess based on existing theory or by applying a theory in a novel context and then conduct an experiment to test this hypothesis. However, this research adopts a novel research approach and investigates an area not previously subject to empirical research and as such the formation of hypotheses was problematic. Hypotheses were drawn from the results of prior studies however, at the outset of the research, these were both numerous and tenuous. Focusing too closely on one or a number of these hypotheses would have resulted in an unacceptable risk of failing to discover novel and interesting knowledge.

The outcome of this reasoning was to initially adopt an inductive exploratory approach in order to form robust hypotheses that could form the basis of subsequent deductive research. Further consideration concluded that the subsequent deductive research would be outside the scope of this research as the initial exploratory stage would require significant work and it is this that forms the content of this doctoral thesis.

3.2.4.2 The rationale for a mixed methods approach comprising observation and semi-structured interviews

First, two principles underlying the final design are described and following this the advantages of adopting a mixed approach are argued.

When considering the final design, the first principle was that a non-obtrusive, observational approach is the most reliable method for obtaining insights relating to learners' natural behaviour. This principle affected the choice of methods in a number of ways. For example, a true experimental design is generally the gold standard in terms of reliability and validity however for the present research, the accompanying structural requirements reduce the authenticity of the study environment and this in turn compromises the external validity of the design. As discussed above, this is one reason a true experimental design was not adopted.

The second principle behind the research design was to adopt the most appropriate method to answer each aspect of the research question. This resulted in a mixed methods approach. The research design consists of a combination of technology based observation and semi-structured interviews. In this case, technology based observation means that software is used to observe and record learners' interactions with feedback. Such observation is used to establish learners' absolute preferences across the feedback types. Semi-structured interviews are used to confirm these preferences and to investigate learners' intentions behind their choices.

Mixed approaches such as this one, adopting methods from both the quantitative paradigm and the qualitative paradigm, have been criticised from a philosophical standpoint. Critics argue that, given that the underlying epistemological and ontological foundations differ between the two paradigms, then the two research approaches are essentially researching different things (Pring 2004: 45). From an ontological perspective such critics may argue that a technology based observation, based on an objectivist ontology, is researching facts relating to an objectivist 'real world' external to the minds of the subjects. On the other hand semi-structured interviews, based on a relativist ontology, are researching a reality internal to the minds of the subjects. The argument is that the latter reality is not necessarily an accurate representation of the real world and in addition is likely to differ between subjects. As discussed earlier in the section 'Theoretical framework', this research takes the position that the two realities are, in fact, closely related. This position is supported by Pring (2004: 56) who suggests that the proposition that the objectivist and relativist realities are different and separate fails to consider the close relationship between them.

A further advantage of a mixed mode approach is the concept of 'triangulation'. Triangulation is a metaphor taken from navigation in which several geometric readings are taken to locate the exact position of an object such as a boat or a piece of land. This metaphor is used to argue that a combination of methods is likely to produce more

accurate findings. Bryman and Bell (2003: 291) outline a number of studies which illustrate how qualitative research methods, such as interviews, can be used to validate or refine findings that emerge from quantitative methods such as surveys.

The final design consists of four studies. In studies one, three and four subjects study an item of technology based instruction incorporating multiple choice questions. After responding to a question, subjects are presented with six different feedback items corresponding to the types comprising the taxonomy of TBI feedback identified in the literature review. Subjects also choose whether to view feedback immediately following a response or at the end of the instruction. A software tool delivers the activity and tracks learners' interactions with the feedback items. This series of technology based observations is supported by study two which comprises semi-structured interviews aiming to establish firstly subjects' reasoning for their choices in study one and secondly, factors that affect these choices.

The next section 'Research methods' provides detail on the methods adopted to collect and analyse data in this research.

3.3 RESEARCH METHODS

A research method can be defined as,

"...a technique for collecting data" (Bryman and Bell 2003: 32).

As stated in the previous section, this research adopts a mixed research design consisting of a combination of technology based observations and semi-structured interviews. In the previous section the rationales for the adoption of this design were explained. This section details each of the methods adopted to collect and analyse data.

3.3.1 Summary

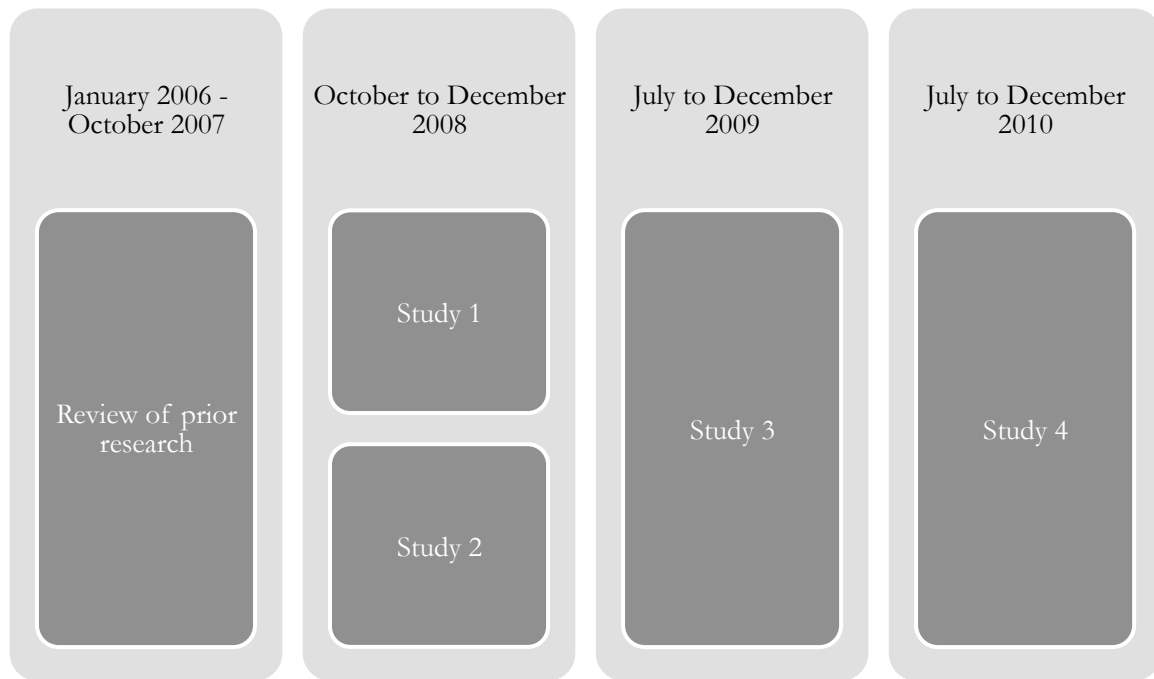
This research investigates learners' preferences for the format of TBI feedback. The term 'format' in this research is defined with reference to the taxonomy of TBI feedback developed in the literature view and refers to a particular selection of attributes from this taxonomy. Learners' preferences are measured by the frequency at which learners choose to view different attributes of TBI feedback when given the choice.

The final design is exploratory in nature and consists of four studies. In studies one, three and four subjects study an item of technology based instruction (TBI) incorporating multiple choice questions. After responding to a question, subjects are presented with six different feedback items corresponding to the six types in the taxonomy of TBI feedback identified in the literature review. Subjects are also given the choice whether to view feedback after responding to a question or after having responded to all the questions. A software tool delivers the activity and tracks each click subjects make in the TBI interface.

Study two comprises a series of semi-structured interviews and aims to identify the subjects' reasoning for their choices in study one and the factors that affect these choices.

The timeline of four studies is shown in the following figure 9. A detailed design of each study is provided in the subsequent chapter 'The four studies –Detailed design, results and analysis'.

Figure 9: The timeline of this research



This section now explains the selection of the research subjects before detailing the methods adopted for collecting and analysing data. The section closes with a consideration of issues relating to reliability and validity.

3.3.2 The selection of the research subjects - Sampling strategies

In research aiming to produce results that can be generalised from a smaller sample population to a larger population, determining a well defined and valid sampling strategy is an important consideration. It is first necessary to determine the larger population.

In this research, who should comprise this larger population was largely determined by the requirement that the studies investigate authentic learning tasks. Once this decision had been made, it was decided to conduct the research at the host institution as it was unlikely that it would be possible to obtain the required access to students and systems at other institutions. The wider population in this case thus comprises students at the host institution.

Two options were considered when identifying the group of learners to be observed. First, forming a group specifically for the purposes of this research. Second, using an intact cohort of students. The former option affords greater control over the research environment however removing learners from the authentic learning context compromises the authenticity of the results due to effects such as the Hawthorne Effect (Bryman and Bell 2003: 45). This refers to the phenomenon of research subjects deviating from their natural behaviour when they are aware that they are participating in a research study. A guiding principle in the design of the present research is that a non-obtrusive, observational approach is the most reliable method for obtaining insights relating to learners' authentic behaviour and thus observing an intact cohort of students was chosen for this investigation.

Once the target population had been restricted to existing cohorts of students at the host institution, the primary restriction became the range of TBI suitable to form the focus of this study. This is explored in more detail in a later section ‘Selection of the items of TBI adopted in this research’ (See page 102).

For the technology based observations adopted in three of the four studies, sampling from within a cohort was unnecessary and data from all subjects was collected. However, sampling was required for study two which comprised semi-structured interviews. The approach taken is detailed in the ‘Design’ section of the reporting of study two (See page 148).

3.3.3 Data collection

For reasons described previously in the chapter ‘research design’, technology-based observation and semi-structured interviews were chosen as the data collection methods for this research.

3.3.3.1 Technology-based observation format

Different approaches to observation were considered including the use of human observers, the use of screen recording software and the development of interface tracking software. The latter approach, termed here ‘technology based observation’ was chosen for two reasons. First, such software would allow for data collection to be automated resulting in a large number of observations thus increasing the validity of the study. Second, as described above, because such an approach is non-intrusive the study is less likely to be compromised by issues such as the Hawthorne effect (Bryman and Bell 2003: 45). This again increases the validity of the research.

The main difficulty in adopting technology based observation was that appropriate software did not exist. Therefore, it was necessary to develop software and this constituted a significant component of the time spent on this research. The development of this tool is described in the section ‘The Development of the Software Tool’ (See page 112).

3.3.3.2 The semi-structured interviews

Detailed descriptions of the approach taken to the semi-structured interviews appear in the ‘Design’ section in the reporting of study two (See page 148).

3.3.4 Data analysis: Quantitative data

This research gathered both quantitative and qualitative data. Different approaches to data analysis were required for each type of data. In this section, approaches taken to the analysis of quantitative data are first discussed.

Statistical techniques can be used to perform a range of tasks. For example they can describe data, infer predictions from data, measure the extent to which different items in a data set vary from each other, measure differences between groups and to establish the relationships between groups. In the present research, quantitative data are collected as the software tool records subjects’ interactions with the multiple choice question interface within the

items of TBI. The base unit of data collected in this research is a 'click'. A click is registered when a subject selects a target item such as a question number, a feedback item or a software setting.

Statistics are used firstly to describe the distribution of these clicks (descriptive statistics) and secondly to establish the significance of certain factors in determining whether subjects elected to view a feedback type (inferential statistics). These are now introduced in turn.

Descriptive statistics are used to describe broad trends occurring in the data. For example to show how frequently subjects click on feedback, show the relative popularity of different feedback types and show simple preferences such as the percentage of subjects who elect to view immediate as opposed to delayed feedback. For the majority of hypotheses in the present research, descriptive statistics in the form of frequency tables, the mean, the median, the mode and the standard deviation along with presentational mechanisms such as histograms are sufficient to demonstrate whether a hypothesis is supported.

However, the present research also hypothesises that certain factors affect whether learners chose to view a feedback type., for example, does a question focusing on the application rather than recall of previously studied material affect whether a learner chooses to view a particular feedback type? And this requires the use of inferential statistics. Other factors considered in the present research include:

- Performance - Whether the learner answers a question correctly or incorrectly
- Gender - The gender of the learner
- Assessment - Whether the materials are subsequently assessed
- Learner level - Learners ability in the subject

This list includes factors that are nominal in nature such as gender, and those on a ratio scale such as learner level. If the research had determined that just one factor affected subjects' choices then a statistical test that compared means would be sufficient to establish the significance of that factor. However, the present research identified multiple factors and thus needed to consider the possible impact of these factors on each other. Given this, two possible tests were considered, linear regression and logistic regression. These are now discussed in turn.

3.3.4.1 Issues in adopting linear regression analysis

Linear regression analysis was initially the preferred test due to the potentially robust nature of any resulting findings. Linear regression is used to produce a model that is able to predict the value of a dependent variable given values for a number of independent predictor variables. In this research, an initial method considered was to produce six linear regression models, one for each feedback type. Factors such as 'question focus', 'performance' and 'gender' would represent the independent predictor variables and the dependent variables would be whether a subject clicked on a TBI feedback type. However, linear regression analysis makes a number of assumptions

regarding the variables that are being analysed and not all these assumptions would hold in this research. This issue is now discussed.

In this research the basic unit of dependent data is nominal with two categories, whether a subject decides to click on a feedback type or not. Linear regression assumes a linear relationship between the dependent variable and one or more of the independent variables and thus the dependent variables are required to be on the interval or ratio scale. In order to satisfy this requirement, a construct on the ratio scale was required and the 'proportion of questions in which a feedback type was selected' was chosen as the unit of measure for this purpose. This measure was given the symbol ' α ' and is described in detail in the following section 'unit of measure'.

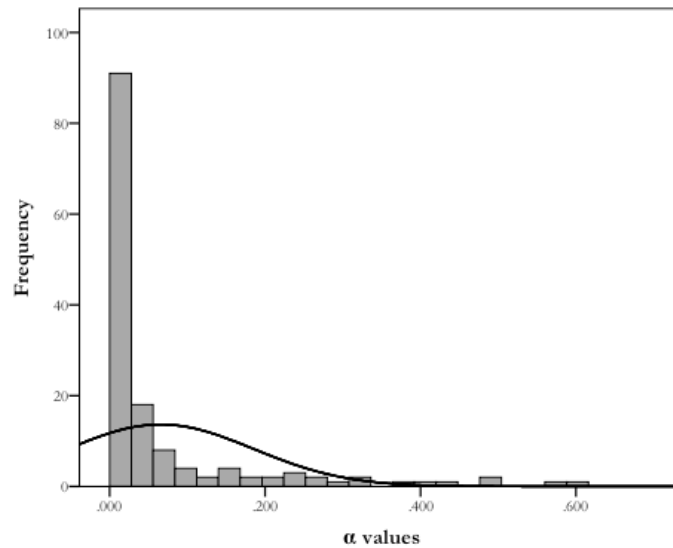
In addition to being on the interval or ratio scale, the dependent variable is also required to be normally distributed. Table 6 shows descriptive statistics relating to the α values for each TBI feedback type. For example the mean α value of 0.9135 for KR type feedback means that, on average, subjects viewed this feedback type for 93.50 per cent of the questions to which they responded. The high levels of skewness and kurtosis show that the variables are far from normally distributed.

Table 6: Descriptive statistics relating to the ' α ' values for each TBI feedback type

	No feedback	KR	KCR	IB	EX	EI1	EI2
N	3551	3551	3551	3551	3551	3551	3551
	0	0	0	0	0	0	0
Mean	0.0726	0.9135	0.2051	0.1246	0.1318	0.062	0.0497
Median	0	1	0.1670	0	0	0	0
Mode	0	1	0	0	0	0	0
Std. Deviation	0.2219	0.2422	0.2377	0.2132	0.2238	0.1656	0.1436
Skewness	3.2170	-2.9210	1.5700	2.4110	2.3550	3.8710	4.4590
Std. Error of Skewness	0.0410	0.0410	0.0410	0.0410	0.0410	0.0410	0.0410
Kurtosis	9.2790	7.3300	2.5320	6.1810	5.6230	16.5540	22.9970
Std. Error of Kurtosis	0.0820	0.0820	0.0820	0.0820	0.0820	0.0820	0.0820

Line three in table 6 shows the mean α values for each TBI feedback type. The distribution of these α values is illustrated in figure 10 in a histogram plotting the frequency distribution of for the IB feedback type across the subjects in this research. The distribution of distribution of α values for other feedback types follows a similar distribution. The values are not normally distributed.

Figure 10: Frequency distribution of α values for the IB feedback



A number of transformations were applied to the α values in an attempt to achieve normally distributed dependent variables. The histograms in the following figures illustrate the results of various transformations applied to data relating to the IB feedback type. None of these transformation, or combinations of these transformations were able to produce data that approximated a normal distribution to an acceptable degree.

Figure 11: Log (left) and Square root (right) transformations

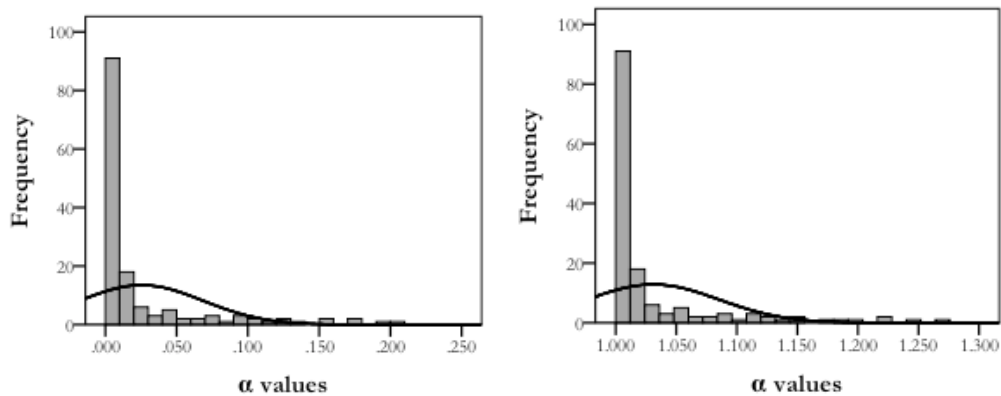
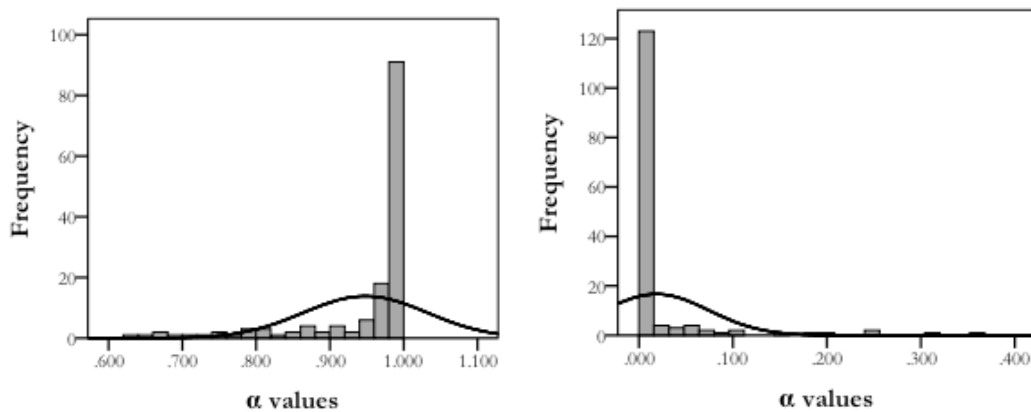
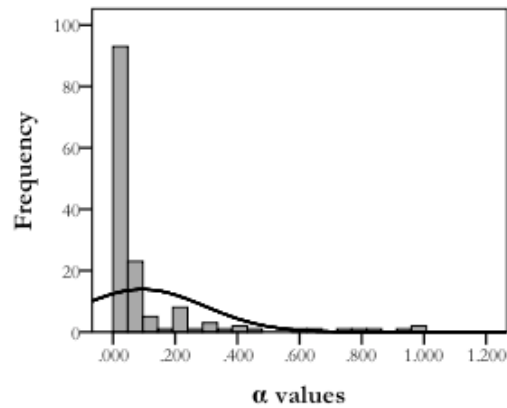


Figure 12: Inverse (left) and Square transformations



An additional measure considered was to focus on data that related to incorrect responses. However, as shown in figure 13, this resulted in data which following a similar distribution.

Figure 13: Data relating to incorrect responses



The inability to determine dependent variables which followed an acceptable approximation to a normal distribution resulted in linear regression being abandoned as an appropriate statistical test in this research.

3.3.4.2 *Logistic regression*

Logistic regression is similar to linear regression in that it attempts to produce a model that predicts a dependent variable from given values of a number of independent predictor variables. However logistic regression is used when the dependent variable takes a binary form. Rather than predicting a value for a dependent variable, a logistic regression model predicts the probability of the dependent variable falling into one of a number of categories.

Thus a logistic regression model can be used in this research to predict the probability that a subject will click on a particular feedback type given particular values for the factors considered in the present research. Logistic regression analysis can also be used to determine whether the factors are significant and the extent to which they affect whether subjects choose to view a feedback type. Critically, logistic regression is appropriate whether the independent predictor variables take either nominal and/or ratio scale form.

An assumption made when performing a standard logistic regression analysis is that all records in the data are independent from each other. In the data structure described below, there are multiple records for each student and thus this assumption is invalid. As a result, a multilevel logistic regression analysis was adopted for the present research to address this issue.

All the assumptions required to conduct such a logistic regression analysis were satisfied in this research and this method was chosen as the statistical test to analyse factors that potentially affect whether subjects chose to view a feedback type.

3.3.4.3 Logistic regression: Organization of data

Seven different logistic regression analyses are performed. One for each of the six feedback types plus one for the condition 'no feedback' in which a subject does not select any of the feedback types after responding to a question. This analysis requires data to be organized at the question level. Each record relates to one student answering one question and contains some or all of the following data:

a) Information relating to the subject:

- The subject's e-mail address
- The subject's gender
- The subject's programme of study
- The year the subject commenced their degree programme
- The subject's score in a course assessment
- The subject's score in their degree programme

b) Information relating the question to which the subject responds

- The test to which the question belongs
- The number of the question within a test
- Whether the material being studied was subsequently assessed
- Whether the subject matter of the test is quantitative or qualitative
- Whether the question tests recall or application of the material being studied
- Whether the question was answered correctly
- Whether the subject selected each feedback type (or did not view feedback)

3.3.4.4 Logistic regression: Procedure

Prior to performing the analyses, variance inflation factors (VIFs) were calculated to identify the presence of any collinearity between the independent predictor variables. Then the logistic regression analyses were performed in two stages. First univariate analyses were performed for each factor to assess the impact of the factor without considering the impact between the factors themselves. Then, multivariate logistic regression analyses were performed that take into account such an impact.

3.3.4.5 Logistic regression: Interpreting the results

The key result of logistic regression is termed the 'odds ratio'. For the analyses in this research, the odds ratio indicates the extent to which a factor affects whether subjects choose to view a feedback type. It is calculated according to the following formula:

$$OR = \frac{\text{Odds of an event occurring in one category in the independent predictor variable}}{\text{Odds of an event occurring in the other category in the independent predictor variable}}$$

Where odds is calculated as follows:

$$\begin{aligned} Odds &= \frac{\text{The probability of the event occurring}}{\text{The probability of the event not occurring}} \\ &= \frac{P(\text{Event})}{P(\text{No event})} \end{aligned}$$

Thus if the odds of an event are zero then there is no chance that the event will occur and the further the odds move away from zero then the more likely it is to occur.

Given this, if the odds ratio equals 1 then the predictor variable has no effect on the likelihood of the event occurring. If the odds ratio is less than 1 then the change in the predictor variable makes the event less likely to happen and the closer to zero the stronger the effect. Conversely if the odds ratio is greater than 1 then the change in the predictor variable makes the event more likely to happen.

3.3.4.6 *Analysis of quantitative data: Units of measure*

The purpose of this study is to determine learner preferences for the format of TBI feedback provided following responses to multiple choice questions. To measure preference we need to determine a construct or unit of measure to represent this concept. The software used in this study records each click a user makes on a feedback item and it is these clicks that must form the basis of any measure. Preference is thus measured by the frequency at which subjects elect to view different feedback attributes. Note that the selection of attributes is not mutually exclusive in this research, subjects are able to select more than one attribute for each question.

Units of measure are required first to capture the frequency at which subjects chose to view each feedback type and also to capture the frequency at which subjects elect to view immediate and delayed feedback. These different measures are now described.

Units of measure to capture subjects' preference for a feedback type for use in the descriptive statistics

This unit of measure could be set at different levels. For example, at the activity level by measuring the total number of clicks on a TBI feedback type over the course of an activity or measuring the number of questions in which a feedback type was selected over the course of an activity. Alternatively at the question level, measuring the number of clicks made on each feedback item for each question in the activity. The prime consideration when selecting the item measure was that the measure was reliable, could be readily interpreted and was suitable for the statistical analyses adopted in this study. Three measures were considered;

Total clicks on a feedback type across the activity: This measure has the advantage that it incorporates all the data captured, it includes every click a subject made. The disadvantage is that repeated clicks of a

feedback type on one particular question would skew results. In addition, instances in which students did not complete all questions would need to be removed from analysis as this would also skew the data. For example, if a subject attempted just one question in a pass through the activity then the number of clicks on a particular feedback type would be recorded as low but this should not be interpreted as the subject having a low preference for that feedback type.

Number of questions in which a feedback type was selected: This measure is reliable, easy to understand and suitable for statistical analysis. The disadvantage is that the measure does not capture how many times an item of feedback is viewed and thus some of the raw data are lost. In addition, as per the previous measure, instances in which students did not complete all questions would need to be removed from analysis.

Proportion of questions in which a feedback type was selected: This latter problem could be mitigated by adopting the proportion of questions in which a feedback type was selected as the unit of measure. This would allow partially completed exercises to be incorporated into the analysis but adds an additional layer of complexity which may be problematic for more advanced analysis.

The latter measure was chosen as this met the stated requirements and was considered to most accurately represent learner preferences. In accordance with the regulations at my host institution, this number is represented as a percentage accurate to four decimal points. This measure is represented in this research by the symbol, α .

Units of measure to capture the frequency at which subjects chose to view each feedback type use in the logistic regression analysis

In the logistic regression analysis, each record in the data relates to one student answering one question. In this case the act of a subject clicking to view a feedback type for a question is considered to represent a preference for that feedback type.

Unit of measure to capture learners' preference for the timing of feedback

In the items of TBI adopted in this research, subjects choose whether to receive feedback after each question or after having answered all the questions. Subjects initially make this choice prior to starting the test but could subsequently change this choice while they take the activity. Two measures were used to capture subjects' preference regarding the timing of feedback. First, subjects' initial choice and second, their position at the end of the activity.

3.3.4.7 Analysis of quantitative data: Use of statistical software

Given the high volume of data and large number of statistical tests required for this research, I decided to use statistical software to assist in the analysis. Three software programs were considered, SPSS, Stata and Gauss.

SPSS was initially chosen due to recommendations from colleagues who proposed that SPSS contained the most efficient user interface and also because of its prominence in the social sciences. It was recognized that the Stata

and Gauss software packages contained more advanced functionality, particularly relating the presentation of results, however this functionality was not initially considered to be required in this research.

SPSS was therefore initially adopted as the statistical package to be used. However, as described earlier in this section, a proposed linear regression analysis was replaced by a multilevel logistic regression analysis. At this point it became necessary to use the Stata package as SPSS did not contain the functionality to perform multilevel logistic regression analysis.

Therefore, in this research, SPSS was used to conduct the descriptive analyses and Stata was used to conduct the multilevel logistic regression.

3.3.4.8 Analysis of quantitative data: Data processing

A large volume of data for a study of this type was gathered. Over three years, data were collected relating to 24 cohorts of students studying one or more of three items of TBI materials each containing between 15 and 50 questions each of which contained six feedback options. Following data processing, data relating to 61,929 responses to questions formed the data set in this research. Users could elect to revisit feedback at any time and could also choose to view feedback either after each question or upon completion of the whole material. Subjects were also allowed to repeat the material.

The raw data captured by the software system required conversion into a format suitable for the statistical tests described above. This data processing was complex and frequently requiring manual intervention on a record by record basis. This process thus consumed a significant proportion of the time spent on the data analysis phase of this research.

This process was exacerbated by the exploratory nature of the research. As new hypotheses were formed, new fields were required to be added to records in the data set.

Data processing primarily comprised the removal of multiple submissions, removing errors from records and adding data required to test hypotheses that emerged during the research. Each of these is now described.

The removal of multiple submissions

An initial analysis of the data revealed that subjects made multiple attempts at the tests. Of these multiple attempts, a large number appeared to be false starts in which subjects would initially open a test, answer one or two questions, then close the test and return later to complete the exercise in full.

In order to remove bias in the data resulting from multiple submissions, including such false starts, I decided to include one submission from each subject. The procedure was performed using the following heuristic: If the e-mail address and test name are the same for two submissions then discard the submission with the fewest questions answered. If the subject answered the same number of questions in more than one submission then discard most recent sessions.

Data identifying individual students were not captured for one item of TBI, the Pre-study Mathematics course described in a later section, in 2009 due to an error in the reporting. Therefore all submissions to the Pre-study Mathematics tests are included in the data set adopted in study three.

Removal of errors

This process of data formatting was complicated by the fact that students firstly made a significant number of errors during while registering to take a test. These including spelling errors, selecting the wrong year, selecting an incorrect degree programme or using a personal email address as opposed to the institutional e-mail address as requested. Before data analysis could commence it was necessary to manually edit the data to remove any obvious errors. Virtually all these errors were resolved by cross referencing each record with data sets held on student cohorts by the host institution. If an error could not be resolved then the record was removed from the data set.

The addition of data

Given the exploratory nature of the research it was not possible to determine in advance all of the data fields that would be required in subsequent analyses. As an example, factor 'learner level' was not an identified factor at the time the software tool was designed. Such data were added to the existing data at a later stage. Further examples of such data include whether material was subsequently assessed, whether a question tests recall or application of the material and whether a subject is quantitative or qualitative.

Much of this addition of data was too complex to be automated and was performed manually. Often, this required each individual student entry to be matched against up to four different data sources in order to obtain the required data. This time consuming process was repeated on each time data were extracted from the database and comprised a significant proportion of the analysis component to the present research.

Care was taken to ensure this process was rigorous and only student records that could be identified precisely against all required sources were carried into subsequent analysis. This rigour did result in a loss of approximately twelve per cent of records.

3.3.4.9 Analysis of qualitative data: Approaches considered

In this research qualitative data are collected from the semi-structured interviews in study two. The purpose of these interviews is to investigate subjects' reasoning for their choices in study one and the factors that affect whether subjects chose to view, or not to view, different TBI feedback types.

A fundamental issue faced when analysing qualitative data is the large volume of language that is produced. A significant role of data analysis strategies is to reduce the quantity of data through techniques such as summarising, discovering patterns and proving or disproving hypotheses. Three strategies for qualitative data analysis were considered:

Pattern matching

In this approach, an expected pattern of results is determined based on a study of prior research (Yin (2002: 136). The raw qualitative data are then analysed to establish the degree to which the data agree with this pattern. This type of research generally attempts to prove causality between a set of circumstances and a set of outcomes. The researcher hypothesises that a set of circumstances will lead to particular outcomes and if these outcomes are confirmed by the data analysis then the research findings give strong support for the predicted causality.

For the present research such an approach would enable the findings of literature review to be confirmed or contradicted. The obvious disadvantage is that any additional factors in the data are not discovered by a pure pattern matching approach.

Grounded theory

Grounded theory, popularised by Glaser and Strauss (1967) takes a different, inductive approach to forming explanations. The core premise is that explanations emerge from the data rather than being hypothesised beforehand. A characteristic of grounded theory is that it consists of a systematic and defined set of procedures. In grounded theory, the researcher starts data analysis with general questions rather than hypotheses.

The advantages of grounded theory are that the approach takes into account the complexity of social situations and provides a strong link between data and theory that can enhance the validity of any findings. In addition, grounded theory provides a systematic approach to data analysis rather than just a set of guided principles. There are a number of criticisms of grounded theory at a conceptual level (Bryman and Bell 2003: 432) but the main drawback for the present research is more practical. Given the relatively small volume of qualitative data in this research, the complex and time consuming process grounded theory requires is not justified.

Explanation building (Analytic induction)

This third approach is similar to pattern matching but incorporates inductive techniques as per grounded theory (Bryman and Bell 2003: 426). The researcher starts with an initial proposition based on prior research and tests this proposition by analysing an initial case. This first stage analysis is used to refine the proposition if required. The revised proposition is then tested against further cases and, again, refined if necessary.

As has been noted, there is no single correct way to conduct qualitative data analysis (Gibbs 2002; Cohen, Manion et al. 2007: 461) and, grounded theory apart, common strategies provide researchers with a set of guiding principles rather than a systematic approach. The underlying principle adopted here when determining a strategy was ‘fitness for purpose’ and given this, explanation building was chosen as the data analysis strategy for the qualitative interviews. The primary reason being that it comprises both deductive and inductive approaches. The literature

review identified a number of factors that can form the basis of deductive enquiry however it is the relative paucity of literature in this specific area that forms the gap this research seeks to address. As such, it was necessary to adopt an approach in which possible further factors were able to emerge during the analysis and this required an inductive element.

3.3.4.10 Analysis of qualitative data: The explanation building approach adopted

The specific approach taken is that described by Yin (2002: 143),

1. Make an initial theoretical statement or an initial proposition
2. Compare this against the findings of an initial case
3. Revise the statement or proposition to match the initial case if required
4. Compare the revised statement or proposition against remaining details in initial case
5. Compare the revised statement against the findings of the remaining cases
6. Repeat this process as required

3.3.4.11 Analysis of qualitative data: Selection of CAQDAS software

The use of CAQDAS software was chosen to assist with the analysis of interview transcripts. CAQDAS software is used to record the process of analysis and provide tools that enable features and relationships in the data to be examined more readily. These advantages are not likely to be as significant in a relatively small data set such as that in the present research, however it was felt that gaining experience in using such tools would be a useful part of my training as a researcher. NVivo was selected as the CAQDAS used.

3.3.5 Reliability and validity

Two criteria for evaluating a research design are reliability and validity. Issues of reliability and validity are discussed throughout this thesis however this section aims to summarise the most salient points. Reliability and validity are addressed in turn.

3.3.5.1 Reliability

Reliability is concerned with the degree to which the findings from a piece of research are repeatable. In other words, is a research design likely to produce consistent results? Bryman and Bell (2003: 74-77) propose three main factors affecting the reliability of a research method,

Stability:

Is a measure stable over time? For example, testing undergraduate mathematics students' knowledge of mathematics one, six and ten years following their graduation is not likely to produce consistent results.

Internal reliability:	Does the testing instrument measure the same concepts across different groups. For example, a questionnaire which asks students to evaluate their level of expertise in mathematics is likely to be interpreted differently when delivered to secondary students and doctoral students.
Inter-observer consistency:	This occurs is a research design is likely to produce different results for each researcher. Addressing this issue is of particular importance in research that requires a great deal of interpretation by the researcher.

A strength of the present research is that the technology based observation approach is largely unaffected by these concerns. One valuable contribution made by this research is that it produces a framework for future research. The taxonomy of TBI feedback, the data collection method used to collect in quantitative data and the approach to gathering qualitative data are replicable.

3.3.5.2 Validity: Considerations relating to the technology based observations

Validity is primary concerned with the extent to which the conclusions reached within a piece of research are actually supported by the evidence (Shadish, Cook et al. 2002: 34). In methods such as technology based observations, validity can be considered from a number of different aspects and four such considerations are introduced here. First construct validity; this considers the degree to which the measures adopted in the research measure what they purport to. Second, internal validity which considers the extent to which the conclusions or explanations purported by the research are actually supported by the data. Third, external validity which considers the degree to which causal relationships established by the research can be generalized to a wider population. Finally, ecological validity which is related to external validity but is specifically concerned with whether any causal relationships established by the research equally apply to “people’s everyday, social settings ” (Bryman and Bell 2003: 34). These four forms of validity are now discussed in more detail.

Construct validity considers the degree to which the measures adopted in the research measure what they purport to (Cohen, Manion et al. 2007: 138). For the technology based observations the primary considerations are first, whether the taxonomy of the feedback types is comprehensive and the feedback types are distinct. Second, the extent to which the examples of feedback presented in the three items of TBI accurately reflect the feedback types. Third, whether subjects’ choices regarding feedback relate to the perceived quality of the feedback and not the format. It is pertinent here that this latter concern is present in any research of this nature.

To address the first consideration, this research draws upon precedents established in within prior research in order to produce the taxonomy. This process is described in more detail in the chapter, ‘Literature review’ (See page 42). The second and third considerations are addressed by adopting a systematic approach to the development of feedback for the three items of TBI (See page 105). This process aims to ensure that feedback is of a consistent format across the items of TBI and is also replicable in future research.

Related to the second and third considerations is to question whether the act of clicking on a feedback type is in fact due to a subject wishing to view that feedback type. There may be additional explanations of the action. For example, subjects may click on a feedback type merely because it has been presented to them. Or this choice may be influenced by the feedback types being presented in a particular order. The former is unlikely as users did not click on all the feedback types all of the time. The latter consideration may be pertinent but is unavoidable. As described in the section, 'The Development of the Software Tool', any arrangement of feedback types contains an inherent order (See page 117).

Internal validity considers the extent to which and claims of causality purported by the research are actually supported by the data. (Bryman and Bell 2003: 34). The analysis of data resulting from the technology based observations comprises two components, first the use of descriptive statistics and second, the use of logistic regression. The descriptive statistics do not aim to establish causality between variables but aim to provide insights on the situation contained in the data. In this research, the validity of conclusions reached via the descriptive statistics is strengthened by the fact that data are gathered repeatedly, from 24 different cohorts. (Although the themes outlined in the 'Discussion' chapter (See page 249) relate to an analysis of data relating to 17 of these cohorts).

On the other hand, logistic regression analysis aims to establish causality. In other words, if the independent variables are changed then a consistent resulting change in the dependent variables will be observed. As described in the section 'Research design', the classic experimental design achieves a high degree of internal validity through the use of two equivalent subject groups to which research subjects are randomly assigned, the use of tests prior and subsequent to the change in the independent variable and through controlling for any identified possible confounding factors. Such measures were not possible in the present research which observes the behaviour of students in authentic educational situations. As such the results of the logistic regression cannot claim a high degree of internal validity.

Despite this, efforts were made as regards the internal validity of the approach. The items of TBI and cohorts of students were chosen to ensure that a reasonable balance of data were captured for each independent variable. In addition, the logistic regression incorporated data from as many variables as possible in order to identify confounding factors. This process is described in more detail in the chapter 'The four studies –Detailed design, results and analysis' (See page 125). Possible confounding factors are discussed in detail in the section 'Reliability and validity' in the report of each study

External validity considers the extent to which results can be generalised to populations other than that of the research subjects (Cohen, Manion et al. 2007: 136). This study took place at one institution, a business school, and it could be hypothesised that learners' preferences for the format of feedback may change for different student populations due to factors such as the subject matter and the particular context in which the TBI material is delivered. This places limits on the generalizability of this research and this is discussed in more detail in the section 'Conclusions: Limitations' (See page 271).

Finally, **ecological validity** is concerned with whether any relationships established by research apply to “people’s everyday, social settings” (Bryman and Bell 2003: 34). In quantitative research within the social sciences, research can be conducted within highly controlled environments that do not reflect the actual setting in which results are to be generalized. However, the technology based observations in the present research occur within an authentic educational context and ecological validity is not a significant concern.

3.3.5.3 *Validity: Considerations relating to the qualitative interviews*

Maxwell (1992: 285-295) reports five different kinds of validity that apply to qualitative research.

Descriptive validity:	Does the research accurately report the statements made by the research subjects?
Interpretive validity:	Does the research accurately capture the "intention, cognition, affect, belief, evaluation and anything else that could be encompassed by what is broadly termed the 'participants' perspective'" integral to the statements made by the research subjects?
Theoretical validity:	How valid are the conclusions reached by the research as a theory of the phenomenon explored? In Maxwell’s view, theories comprise two components. First, the concepts or categories that are the object of the theory and second, the stated relationships between them. Theoretical validity therefore considers both the validity of the concepts or categories as constructs and the validity of the stated relationship.
Generalizability:	This considers the extent to which results can be generalized to populations other than the research subjects of the study.
Evaluative validity	This applies to research in which the researcher applies an evaluative framework to the objects of research. For example, when a researcher makes judgments regarding the behaviour of the research subjects. Evaluative validity considers the validity of the evaluative framework.

In this research, consideration was made throughout the design and conduction of study two to address the issues of descriptive, interpretive and theoretical validity. As described in detail in report of “study two’ (See page 150) potential threats to internal validity were considered in relation to the three participating agents; the testing instrument, the research subjects and the interviewer. Care was taken to ensure that the testing instrument did elicit the intended information and that the questions did not steer subjects towards a particular response. Second, to avoid bias in subjects’ responses due to interference from external events, the interviews were conducted over two consecutive days shortly after subjects completed their course and prior to their exam. Interviewer attribute effects were mitigated by informing subjects that the interviews would be anonymous.

Caution needs to be applied when attempting to generalize the findings of the qualitative interviews. The interviews were conducted with ten subjects from one cohort of students and relate to one item of TBI. The primary purpose of the interviews within the present research was to provide insight into subjects' behaviour during the initial technology based observation and to reveal potential avenues for further investigation.

3.3.5.4 Reliability and validity: Triangulation

One further consideration relating to reliability and validity is triangulation. This refers to the use of two or more methods of data collection and the use of triangulation can give additional support to any claims that are made. In this research, the qualitative statements made by subjects in study two are used to confirm or contradict conclusions reached regarding the quantitative data collected in study one. This is discussed previously in this thesis in the section 'The final research design' (See page 80).

3.4 KEY ASSUMPTIONS

As described in the section 'Theoretical framework' (See page 65) the interpretation of results in this research is premised upon constructivist assumptions regarding the nature of reality and knowledge, the nature of the learner and the nature of learning. In addition, assumptions are made regarding the internal validity of the research findings and these are discussed above under "Reliability and validity".

A further key assumption in the design of this research is that the frequency at which subjects select an item of feedback reflects a preference for the corresponding attribute in the taxonomy of TBI feedback. This broad assumption has a number of aspects which were introduced briefly above the previous section 'Research methods: Reliability and validity' (See page 96). These aspects are now discussed in more detail.

3.4.1 Assumption one: Learners selection of an item of feedback relates to its type and not the perceived quality

In research of this type it is possible for subjects' interactions with feedback to be affected by the perceived quality of the feedback. For example, if the quality of a particular feedback type is considered poor then the frequency at which subjects elect to view that type may be lower than if the quality was considered to be high.

As discussed in a later section 'The three items of Technology Based Instruction (TBI)' (See page 105), measures were taken to address this issue however the assumption remains.

3.4.2 Assumption two: The frequency at which learners click on items of feedback is not related the manner in which they are presented in the user interface.

It is possible that subjects will click more frequently on items that are displayed more prominently in the user interface. As discussed in the section “The development of the software tool” (See page 117), the user interface was designed to minimize the possibility that subjects would be led to click on particular items of feedback more frequently. For example, the buttons for all feedback types were given an identical design. However all the arrangements considered, including the final design, contain an inherent order and the possibility exists that this had an effect on the frequency at which subjects clicked on the different feedback types.

3.4.3 Assumption three: The items of feedback accurately reflect their corresponding attributes in the taxonomy of TBI feedback

As discussed under ‘Reliability and validity’ (See page 96) in the section ‘Research Methods’, the validity of this research depends upon the extent to which the items of feedback presented in the TBI correspond to the feedback types in the taxonomy of TBI feedback.

The development of the items of feedback is discussed in detail in the section ‘The three items of Technology Based Instruction (TBI)’ (See page 102). The primary strategy was to replicate the format adopted in prior research where possible. In particular formats were chosen to match the definitions of feedback types identified by Dempsey and Litchfield (1993) and Kulhavy and Stock (1989), the two studies from which the majority of feedback types in the taxonomy are taken.

3.5 THE THREE ITEMS OF TECHNOLOGY BASED INSTRUCTION (TBI)

This section provides information relating to the three items of TBI used in this research. The section starts with an explanation as to why these three items were chosen.

3.5.1 Selection of the items of TBI adopted in this research.

The initial intention in this research was to adopt one item of TBI and then investigate learners' preferences for the format of feedback in this TBI in a range of different conditions.

At the outset of this research, the 'Ethics and Strategy exercise' (described below) was chosen to be this one item of TBI, primarily due to considerations relating to access. A further reason was pragmatic, the majority of the feedback types required already existed for this exercise. The Ethics and Strategy exercise formed the basis of studies one and two however this was removed from the curriculum in 2009 and alternatives were required. Four alternative items of TBI were in use at the host institution that incorporated the use of multiple choice questions:

- The Accounting Primer course
- The Pre-study Mathematics course
- An introductory course in finance
- A course on project management

From these four, the 'Accounting Primer course' and the 'Pre-study Mathematics course' were selected. The rationale being that subjects in study two had stated that subject matter, whether a material tests qualitative or quantitative material, would affect whether they chose to view a feedback type. The Pre-study Mathematics course was chosen as an example of a purely quantitative subject and the Accounting Primer course was chosen as an example of a subject involving both qualitative and quantitative topics. These subjects contrasted with the largely qualitative nature of the Ethics and Strategy exercise adopted in study one.

These two courses had the additional advantage that all students entering the host institution were required to complete one or both of them and this enabled a large quantity of data to be collected.

3.5.2 Description of the items of TBI adopted in this research.

The content and format of each item of TBI are now described.

3.5.2.1 The Ethics and Strategy exercise

This exercise covers two topics from a 20-hour taught course on Business Informatics, 'Developing an Information Systems Strategy' and 'Managing Ethical Issues in Information Systems'. The exercise comprises three hours of self-study and one of the 20 class-based taught hours.

The exercise contained two components, online presentational materials in the form of scanned extracts from the course textbook plus a multiple choice question based activity. Subjects were asked to study the presentational materials in advance while the multiple choice question based activity was delivered in class.

50 questions were incorporated into the multiple choice test. These were selected from a larger bank of questions comprising questions written by the course lecturer, questions written by a previous lecturer and questions developed by an educational publisher. The questions adopted for the online activity were selected from this question bank by the course lecturer using the criteria of relevance and quality.

3.5.2.2 The Accounting Primer course

The Accounting Primer is a 10-hour online course introducing key concepts in accounting. This course is made available to subjects for the two months immediately prior to their arrival at the host institution. Subjects study the course at a time and place of their choosing.

The material is split into sections and a number of multiple choice questions appear at the end of each section. Data related to questions from sections 'the Balance Sheet' (6 questions), the 'Profit and Loss Statement' (7 questions) and the 'Cash Flow Statement' (5 questions) were analysed in this research.

A screenshot from the course is presented below. The different sections of the course can be seen in the menu on the left hand side.

Figure 14: A screenshot showing the structure of the Accounting Primer course

The screenshot displays the 'Accounting Primer' course interface. On the left, a sidebar menu lists the course structure: Orientation (4), Balance Sheet (12), Income Statement (12), Cash Flow Statement (6), and Course Review (4). The main content area is titled 'Balance Sheet Formats' and explains that balance sheets can use different layouts and classifications. It introduces two common formats: horizontal and vertical.

Horizontal (represented by a hamburger menu icon) is commonly used in the US. Assets are listed in decreasing order of liquidity. The Accounting Identity is expressed as:

$$\begin{array}{rcl} \text{Current Assets} & + & \text{Current Liabilities} \\ + & & + \\ \text{Fixed Assets} & = & \text{Long Term Debt} \\ & & + \\ & & \text{Equity Capital} \end{array}$$

Vertical (represented by three vertical bars) is commonly used in the UK. It is the traditional method of displaying a balance sheet. The Accounting Identity is expressed as:

$$\begin{array}{rcl} \text{Assets} & - & \text{Liabilities} = \text{Equity Capital} \\ \text{Fixed} & & \text{Short} \\ + & & \text{Term} \\ \text{Current} & & + \\ & & \text{Long} \\ & & \text{Term} \end{array}$$

Both sections include links to 'click here to work through an interactive sheet' and 'Other examples: web version, Excel version'. The bottom of the interface features a navigation bar with tabs for Orientation, Balance Sheet, Income Statement, Cash Flow Statement, and Course Review, along with a copyright notice for Imperial College Business School.

3.5.2.3 The Pre-study Mathematics course

The Pre-study Mathematics course is a 12-hour course similar in format to the Accounting Primer course but covering key concepts in mathematics. Again, this course is made available to subjects for the two months immediately prior to their arrival at the host institution. Subjects study the course at a time and place of their choosing.

As per the Accounting Primer course, the material is split into sections and a number of multiple choice questions appear at the end of each section. Data related to questions from sections 'Basic Maths One' (5 questions), 'Basic Maths Two' (6 questions), 'Differentiation and Optimization' (4 questions), 'Integration' (4 questions) and 'Sequences and Financial Mathematics' (6 questions) were analysed in this research.

A screenshot from the course is presented below. Again, the different sections of the course can be seen in the menu on the left hand side.

Figure 15: A screenshot showing the structure of the Pre-study Mathematics course

Weekend EMBA: Pre-study Mathematics (2009-2010)

Build Teach Student View

Table of Contents for Pre-Study Mathematics

- 1 Introduction
- 2 Basic Maths 1: Algebra, Equat
- 3 Basic Maths 2: More Algebra,
- 4 Differentiation and Optimizat
- 5 Integration
- 6 Sequences and Financial Mat
- 7 Course Summary

Your location: [Home Page](#) > [Pre-Study Mathematics](#) > [Integration](#) > [Study Pages](#) > [Areas Under Curves](#)

Pre-Study Mathematics - Areas Under Curves

Learning Objectives:

To introduce one of the two main applications of Integration in the business context: Finding the area under a curve.

Instructions:

Read though the text and answer the questions below.

Time: 10 minutes

Definite Integration - Areas under curves

One of the main uses of Integration within the business context is to measure the area under a curve – or the area between a curve describing a function and the x-axis.

Once the three items of TBI had been chosen, it was necessary to process these materials into the format required by the research design. The first part of this process was to change the technical format so that it was compatible with the software tool described in the next section. The second, and more significant, part of this process was to determine and create the feedback types. This latter procedure is now explained.

3.5.3 Determining the feedback types

An initial task was to determine how each feedback attribute in the taxonomy of TBI feedback, shown in table 5 above (See page 46), would be represented in the three items of TBI.

3.5.3.1 Representation of the attributes in the taxonomy of TBI feedback

When determining how to represent the feedback attributes, the primary strategy was to replicate those adopted by prior research where possible. Formats were chosen to match the definitions of feedback types identified by Dempsey and Litchfield (1993) and Kulhavy and Stock (1989), the two studies from which the majority of feedback types in the taxonomy are taken.

The 'No feedback (NF)', 'Knowledge of result (KR)' and 'Knowledge of correct response (KCR)' feedback types are self-explanatory and no decision was required as to the format they would take. They are represented as follows:

No feedback (NF):	The subject does not view any of the feedback types
Knowledge of result (KR):	A statement as to whether a subject's response is correct or incorrect
Knowledge of correct response (KCR):	The presentation of the correct response

Representations of the explanatory (EX), instruction based (IB) and extra-instructional (EI) feedback types required greater consideration. The forms adopted are now explained in turn.

3.5.3.2 Representation of Explanatory (EX) type feedback

In the Ethics and Strategy exercise, explanatory (EX) type feedback comprises explanations as to why each answer is correct or incorrect as shown in the example below from the Ethics and Strategy exercise.

Figure 16: Example of EX type feedback from the Ethics and Strategy exercise.

The use of computers to combine data from multiple sources and create electronic dossiers of detailed information on individuals is called:

☐ profiling.
☐ spyware.
☐ spamming.
☒ targeting.

▼ Further explanation:

'Profiling' is the correct answer.

'Spyware' - "is a specific term and is not directly related to 'profiling'.

'Spamming.' - "is the abuse of electronic messaging systems to indiscriminately send unsolicited bulk messages." (Wikipedia 14/10/08)

'Targeting' - is partially correct but not a precise answer.

The same format of explanatory (EX) type feedback was used for the Accounting Primer and Pre-study Mathematics courses where appropriate. Where necessary, an additional format was used which comprised a presentation of the process required to reach the correct answer.

Figure 17: Example of (EX) type feedback from the Accounting Primer course

▼ Further explanation

Completed

1.
If Sales Revenue = 937, opening stock = 463, purchases = 453 and closing stock = 307, what is the gross profit?

Student Response	Value	Correct Answer
Answer: not answered	0%	328

Score: 0/1
Comments:

Net Interest	
Interest Received	40
Interest Expense	(360)
Profit before Tax	5,400
Taxation	(1,700)
Profit for the financial year	3,700

3.5.3.3 Representation of Instruction Based (IB) feedback

Instruction based (IB) feedback comprises the presentation of the relevant section of the instructional material. The following examples illustrate the format of Instruction based (IB) type feedback for each of the three items of TBI. IB type feedback for the Ethics and Strategy exercise comprised presenting the relevant extract from the digitised textbook.

Figure 18: Example of IB type feedback from the Ethics and Strategy Exercise

▼ Book reference:

Management Information Systems - Laudon & Laudon, Page 131
© 2007, 2006 Pearson Education Inc.

Think of all the ways you generate computer information about yourself—credit card purchases; telephone calls; magazine subscriptions; video rentals; mail-order purchases; banking records; local, state, and federal government records (including court and police records); and visits to Web sites to read Web materials, use search engines, and write blogs (see Chapter 10). Put together and mined properly, this information could reveal not only your credit information but also your driving habits, your tastes, your associations, intended purchases, political views, and interests. What you thought was private, in fact, can quickly become public.

Companies with products to sell purchase relevant information from these sources to help them more finely target their marketing campaigns. Chapters 3

IB type feedback for the Accounting Primer and Pre-study Mathematics courses comprised presenting the relevant screen from the course material.

Figure 19: Example of IB type feedback from the Accounting Primer course

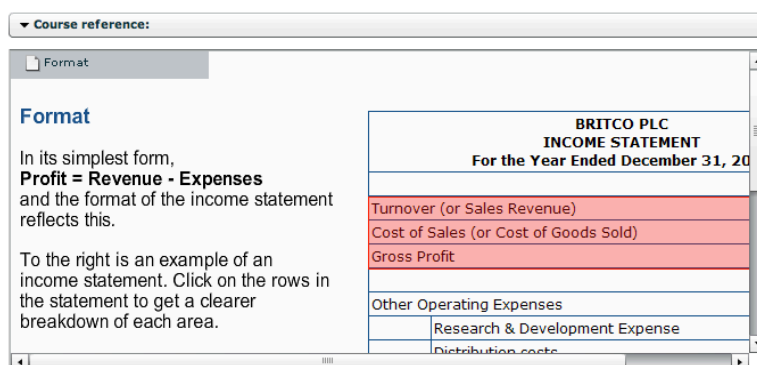
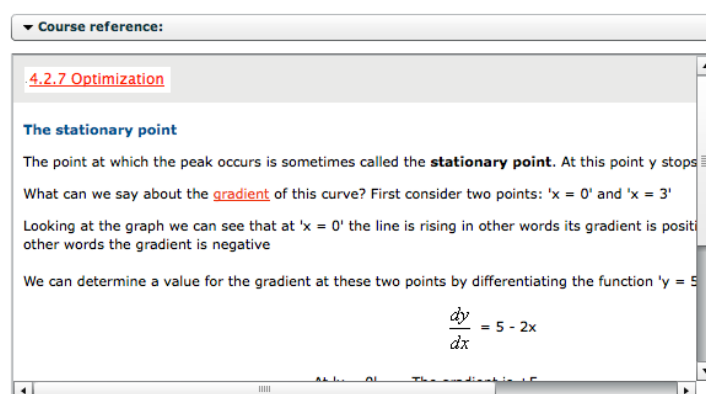


Figure 20: Example of IB type feedback from the Pre-study Mathematics course



3.5.3.4 Representation of Extra-instructional (EI) type feedback

Extra-instructional (EI) type feedback consists of additional information that has not been supplied to subjects previously. This type of feedback can take different formats and thought was required to determine the format adopted in this research. The term 'extra-instructional feedback' was originally defined by Kulhavy and Stock (1989) who propose that such feedback can comprise,

“Examples or analogies not part of the instruction, or when new information is introduced into the feedback message in order to clarify its meaning.” (Kulhavy and Stock 1989: 286)

In order to broaden the relevance of this research, two different types of elaborative feedback were incorporated.

Extra-instructional 1 (EI1): Additional examples of the concept introduced in the question

Extra-instructional 2 (EI2): Definitions and/or explanation of terminology and rare words.

Examples of extra-instructional 1 (EI1) type feedback for the Accounting Primer and Pre-study Mathematics courses are shown in the following screenshots.

Figure 21: Example of EI1 type feedback from the Accounting Primer course

▼ Examples:

The following are examples of how the accruals concept is used to inform the creation of financial statements.

British Land
Net rental income
Rental income is recognised on an accruals basis. A rent adjustment based on open market estimated rental value is recognised from the rent review date in relation to unsettled rent reviews. Where a rent-free period is included in a lease, the rental income foregone is allocated evenly over the period from the date of lease commencement to the earliest termination date.
 (For full notes click [here](#))

Marks and Spencer
Revenue
E. Refunds and loyalty scheme accruals Accruals for sales returns and loyalty scheme redemption are estimated on the basis of historical returns and redemptions and these are recorded so as to allocate them to

Figure 22: Example of EI1 type feedback from the Pre-study Mathematics

▼ Examples:

What is the stationary point for this function: $y = 4x - 3x^2$

Start	$y = 4x - 3x^2$
Differentiating:	$\frac{dy}{dx} = 4 - 6x$
So, the equation we need to solve to find 'x' is:	$0 = 4 - 6x$
Which equals:	$6x = 4$
So:	$x = 2/3$
Now we have 'x', we can calculate 'y' by placing 'x=2/3' into the original formula: $y = 4x - 3x^2$	$y = (4 \times 2/3) - (3 \times (2/3)^2)$
Which equals:	$y = 4/3$
So our stationary point is:	$(0.67, 1.33)$

Examples of extra-instructional 2 (EI2) type feedback for each course are shown in the following screenshots.

Figure 23: Example of EI2 type feedback from the Ethics and Strategy exercise

▼ Explanation of terms:

RFID
 "Radio-frequency identification (RFID) is an automatic identification method, relying on storing and remotely retrieving data using devices called RFID tags or transponders. An RFID tag is an object that can be applied to or incorporated into a product, animal, or person for the purpose of identification and tracking using radio waves." ([Wikipedia](#) 16/10/08)

Biometric
 "a branch of biology which deals with its data statistically and by mathematical analysis" ([YourDictionary](#) 16/10/08)

Figure 24: Example of EI2 type feedback from the Accounting Primer course

▼ Explanation of terms:

Accruals Concept
This states that when dealing with revenues and expenses it is important to reflect rights and obligations to make and receive cash, and not the actual cashflows themselves. Thus for example, revenue on the income statement shows not only cash receipts from customers but also those sales on credit that have been made.

Expenses
Expenses are defined as the cost of resources used up in an operating period. Expenses are not the same thing as cash expenditures.

Figure 25: Example of EI2 type feedback from the Pre-study Mathematics course

▼ Explanation of terms:

Differentiation
Differentiation is one of the main branches of Calculus. This is the process of computing a derivative. The derivative states the rate of change in one variable in relation to another variable.

$$\frac{dy}{dx}$$

This notation indicates the derivative function of a function $y(x)$. An alternative notation is $y'(x)$.

Gradient
The gradient of a line is a measure of its steepness. It is calculated as: Gradient = (Change in y) / (Change in x)

Stationary point
The stationary point is the peak of a graph, it is the point at which the slope of the tangent is zero, i.e. horizontal. At this point y stops rising, and begins to descend.

3.5.3.5 Representation of immediate and delayed feedback

It was also necessary to determine what comprised immediate and delayed feedback. It was logical that immediate feedback should be provided following a subject's response to a question. However, it was necessary to determine the length of the delay for the delayed feedback. In prior research, the length of the delay has varied (See page 45). In this research, delayed feedback is that provided after a learner has responded to all questions in the part of the TBI being studied.

3.5.3.6 Representation of feedback load

Load refers to the quantity of feedback provided. In this research, load, is calculated by the number of feedback types viewed by a subject for a question.

3.5.4 Consideration of 'Quality'

As discussed earlier (See page 97), there is a risk to the internal validity of this series of studies in that a confounding factor may be the perceived quality of the feedback. In other words, subjects' interactions with feedback may relate to their perceived quality of the feedback and not the format. Three steps were taken to mitigate this issue. First, feedback was collated from different sources where possible in order to mitigate issues relating to individual authors. Second, in order to facilitate replication in subsequent studies, feedback was developed in a structured

manner and in a consistent format across all questions. Third, great care was taken to ensure that feedback was relevant and useful.

When it was necessary to develop feedback components specifically for this research, care was taken to follow the principles described above and all work was conducted in conjunction with the course lecturer.

3.5.5 Summary: Representation of the attributes in the taxonomy of TBI

Table 7 below shows how the different elements of feedback within the testing instrument match to different attributes in the taxonomy of TBI feedback.

Table 7: Representation of attributes in the taxonomy of TBI feedback

Attribute	Possible values	Representation in testing instrument	Notation
Verificational	No feedback	The learner does not view any of the feedback types	NF
	Knowledge of result	A statement as to whether a subject's response is correct or incorrect	KR
Elaboration - content	Knowledge of correct response	The presentation of the correct response	KCR
	Explanatory feedback	Explanations as to why each answer is correct or incorrect	EX
	Instruction based	The presentation of the relevant section of the instructional material.	IB
	Extra-instructional	Two types of extra-instructional feedback are provided: 1. Additional examples relevant to the question 2. Explanations of rare words and terminology	EI1 EI2
Elaboration - load	Variable	Encapsulated by total clicks across feedback types	L
Timing	Immediate	Provided following a learner's response to a question.	I
	Delayed	Provided after a learner has responded to all questions in the activity.	D

3.6 THE DEVELOPMENT OF THE SOFTWARE TOOL

In each of the technology based observations, a software tool presents subjects with a series of multiple choice questions comprising part of an online instructional activity. Following each question, subjects are presented with six feedback types corresponding to those in the taxonomy of feedback. The software tool records subjects' interactions with each feedback type. This research was unable to identify any existing software able to perform this task and a software tool was developed specifically for this research.

This work was a significant task and comprised a significant proportion of the time spent on this research. This section first gives an overview of the software tool through a series of screenshots that illustrate the user pathway and the administration interface. Following this the development of the tool is discussed in detail.

3.6.1 Overview

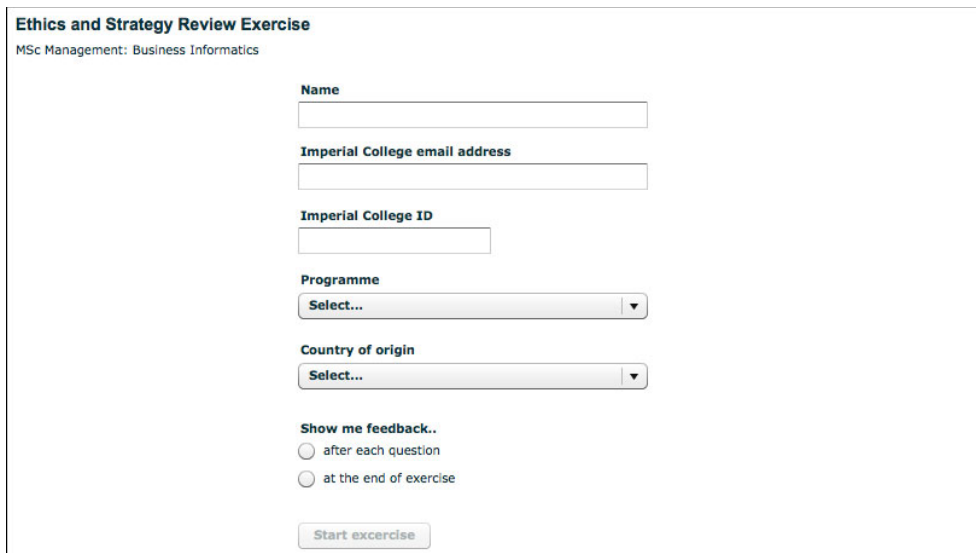
3.6.1.1 *The user pathway*

The following series of screens demonstrate the functionality of the tool from the users' perspective.

Screen 1: The registration screen

First, the software tool captures personal data from the subject. At this stage, subjects also elect to view feedback either after each question or at the end of the exercise.

Figure 26: The registration screen



The screenshot shows a registration form titled "Ethics and Strategy Review Exercise" with the subtitle "MSc Management: Business Informatics". The form contains the following fields and options:

- Name**: A text input field.
- Imperial College email address**: A text input field.
- Imperial College ID**: A text input field.
- Programme**: A dropdown menu with "Select..." as the current selection.
- Country of origin**: A dropdown menu with "Select..." as the current selection.
- Show me feedback..**: Two radio button options: "after each question" and "at the end of exercise".
- Start exercise**: A button at the bottom of the form.

Screen 2: The question screen

The software tool then presents the first multiple choice question to the subject. Each question has four possible answers. At this stage subjects can also amend their choice regarding the timing of feedback and elect to view their current overall score.

Figure 27: The question screen

The screenshot shows a web interface for an "Ethics and Strategy Review Exercise" under the heading "MSc Management: Business Informatics". The main content area displays a question numbered "1": "The introduction of new information technology has a:". Below the question are four radio button options: "dampening effect on the discourse of business ethics.", "ripple effect raising new ethical, social, and political issues.", "beneficial effect for society as a whole, while raising dilemmas for consumers.", and "waterfall effect in raising ever more complex ethical issues.". To the right of the question area, there is a "Show me feedback.." section with two radio buttons: "after each question" (selected) and "at the end of exercise". Below these are navigation buttons (left arrow, "1", right arrow) and a vertical list of question numbers from 1 to 10, with question 1 highlighted in blue. At the bottom right of this section is a "Show my score" button.

Screen 3: The feedback screen

At a time selected by subjects, feedback is presented. There are six feedback options corresponding to the six types in the taxonomy of feedback.

Figure 28: The feedback screen

This screenshot shows the same interface as Figure 27, but with the feedback section expanded. The question and options remain the same. In the feedback section, the "after each question" radio button is still selected. Below the navigation buttons, a vertical list of question numbers from 1 to 10 is shown, with question 1 highlighted in blue. Below this list is a "Show my score" button. The main content area now includes six expandable feedback boxes at the bottom, each with a right-pointing arrow and a label: "Did I answer correctly?", "What was the correct answer?", "Book reference:", "Further explanation:", "Examples:", and "Explanation of terms:". The first box, "Did I answer correctly?", is currently expanded, showing a text input field.

The subject clicks a type to view the feedback and the feedback item is revealed.

Figure 29: The feedback screen showing an example of IB type feedback

Ethics and Strategy Review Exercise
MSc Management: Business Informatics

1

The introduction of new information technology has a:

- ☐ dampening effect on the discourse of business ethics.
- ☒ ripple effect raising new ethical, social, and political issues.
- ☐ beneficial effect for society as a whole, while raising dilemmas for consumers.
- ☐ waterfall effect in raising ever more complex ethical issues.

▶ Did I answer correctly?

▶ What was the correct answer?

▼ Book reference:

Management Information Systems - Laudon & Laudon, Page 129
© 2007, 2006 Pearson Education Inc.

A MODEL FOR THINKING ABOUT ETHICAL, SOCIAL, AND POLITICAL ISSUES

Show me feedback..

☒ after each question

☐ at the end of exercise

◀ 1 ▶

1

2

3

4

5

6

7

8

9

10

Show my score

3.6.1.2 The administration interface

Figure 30 below shows the administration interface. This interface provides an overview of all user records, allows these records to be filtered according to the fields captured in the registration screen and to be edited or deleted.

Figure 30: A screenshot of the reporting interface

MCQ Results 10 students selected [Delete selected results](#)

College ID: Email: Name: Gender: Score between: and

Quiz Name: Course: Country: Test date between: and

☐ Remove duplicate quizzes

Quiz Name	Email	Name	Gender	College ID	Score	Programme	Country	Date / Time
Basic Maths Quiz 1			m		5	Full time MBA		Sun Sep 11 08:52:34
Basic Maths Quiz 1			m		0	Full time MBA		Sun Sep 11 08:49:20
Basic Maths Quiz 1			m		5	Full time MBA		Thu May 10 15:44:46
Basic Maths Quiz 1			f		5	Full time MBA		Sun Jun 17 13:47:06
Basic Maths Quiz 1			f		0	Full time MBA		Mon Jul 4 23:06:34 G
Basic Maths Quiz 1			m		5	Full time MBA		Sun Jan 22 17:42:58
Basic Maths Quiz 1			m		5	Full time MBA		Sat Oct 1 16:47:28 G
Basic Maths Quiz 1			m		1	Full time MBA		Thu Oct 13 13:08:31
Basic Maths Quiz 1			m		5	Full time MBA		Fri Nov 11 20:37:04 G
Basic Maths Quiz 1			m		5	Full time MBA		Wed Jan 11 15:42:40
Basic Maths Quiz 1			m		5	Full time MBA		Sat Oct 8 01:11:38 G
Basic Maths Quiz 1			m		5	Full time MBA		Thu Oct 13 13:10:17
Basic Maths Quiz 1			f		5	Full time MBA		Wed Sep 21 00:49:53

16 records 14 unique emails

[Individual Result Detail](#) [Feedback Reports](#) [Feedback Visualisations](#) [Click Totals](#) [Question Click Totals](#) [Group Totals](#) [Timing Data](#)

[Export Individual Report \(CSV\)](#)

[Export Group Report \(CSV\)](#)

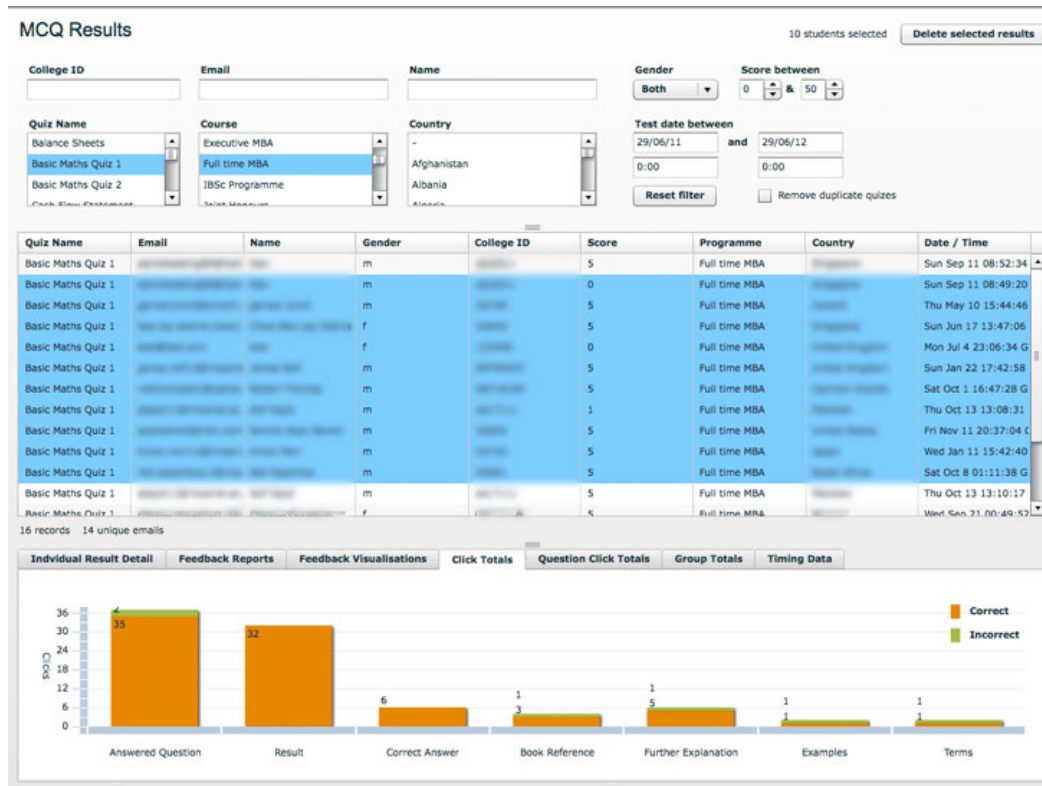
[Feedback Ratios Report \(CSV\) - new](#) <- uses new feedback table

☐ Merge sessions with same email

[Feedback Ratios Report \(CSV\) - original](#) <- max ~100 records

The tracking mechanism tracks each click a user makes in the activity interface. As such a significant volume of data are recorded and the administration interface contains data analysis functionality to provide descriptive views on these data. The lower portion of figure 31 below shows one example of such a view.

Figure 31: A screenshot of the reporting interface



At the initial planning stage, statistical analysis functionality was to be embedded within the software tool to facilitate rapid exploratory analysis. Basic descriptive functionality is incorporated however it was not time efficient to embed the more sophisticated statistical tests into the tool. As a result, functionality was added to the administration interface to enable raw data to be exported in a format that can be subsequently imported into statistics software such as SPSS and Stata. All analysis reported in this thesis was conducted in one of these two software packages.

3.6.2 Software development

3.6.2.1 The development team

The development team comprised myself and one software developer. The volume of work was approximately equal however the nature of tasks differed. I led the development process, conducted the requirements analysis, built prototypes, made key decisions relating to the system architecture, managed the software testing and conducted the deployment. The software developer conducted all the coding appearing in the deployed version.

3.6.2.2 *The development approach*

The approach taken adopted a number of characteristics from “agile software development”. These included an iterative development path, an emphasis on face-to-face meetings over documentation and the use of working prototypes (Agile Alliance 2012). While the approach was iterative, the work performed can be classified into six stages:

- Stage one: The requirements analysis
- Stage two: Specification of detailed requirements
- Stage three: Software design
- Stage four: Implementation
- Stage five: Software testing
- Stage six: Deployment

Each stage is now discussed.

3.6.2.3 *Stage one: The requirements analysis*

The software tool serves three functions. First, it acts as the delivery mechanism for the online multiple choice question based instructional activity. Second, the tool tracks subjects’ interactions with the activity by storing the time and target element of each ‘click’ made by a subject. Third, the tool contains functionality for data analysis . The aim of the requirements analysis performed was to determine broad requirements under each function. These were as follows:

To act as a delivery mechanism, the software tool was required to:

- Be able to import different sets of multiple choice questions and present these to the user.
- Offer the user the choice of immediate and delayed TBI feedback.
- Present the user with the TBI feedback types identified in the literature review following each multiple choice question.
- Be accessible via the Internet so that the multiple choice question based activities can be delivered to different locations.

To track subjects’ interactions, the software tool was required to:

- Capture the following information on a user: name, e-mail address (a unique identifier), course of study, gender and nationality.
- Track all a user’s interactions with the activity. That is, each component a user clicks and the time at which these clicks occur.

For data analysis, the software tool was required to:

- Produce summary reports of data across all or a number of subjects
- Report interactions by an individual subject.
- Filter data according to date, gender and nationality.
- Provide functionality enabling records to be edited and deleted.
- Provide descriptive views of the data.
- Include functionality to enable data to be exported in a format compatible with statistical software packages.

3.6.2.4 *Stage two: Specification of detailed requirements*

In order to capture a detailed functional specification, I produced a ‘wireframe’ prototype comprising diagrammatic representations of each screen. This process established each ‘user pathway’, in other words the different routes each user would take through the system and the logic of the functionality they encounter. This wireframe prototype comprised the documentation for subsequent software design activity.

3.6.2.5 *Stage three: Software design*

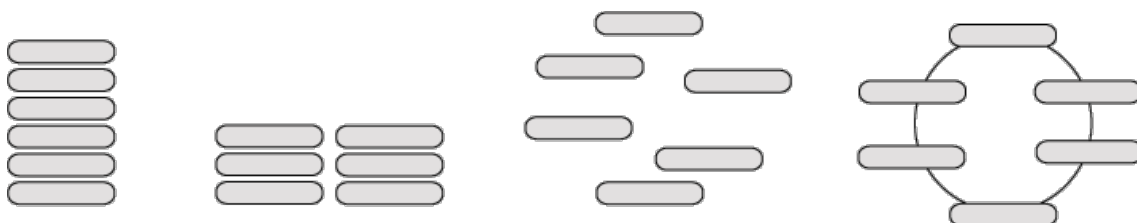
This stage comprised refining the wireframe prototype into the final user interface and determining the system architecture. The latter involved determining the detail of the internal algorithms, the design patterns and the database architecture.

I initially conducted this work myself, however, once I had started work on subsequently coding the system it became clear that I required assistance in order to complete the software tool within the time allocated to this research and I engaged the assistance of a software developer to conduct the coding of the tool. At this stage we refined and amended my initial software design particularly with regard to extensibility, modularity, security and usability.

Note on the presentation of the software types in the user interface:

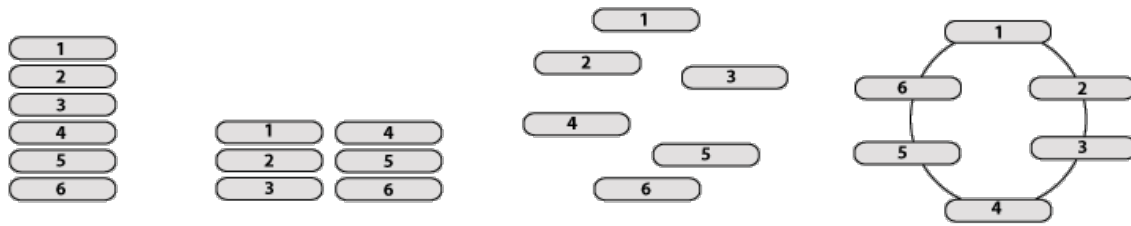
Beyond the general user interface requirements of any web-based educational tool, such as a logical navigation flow and a consistent look and feel, it was necessary that the testing instrument did not lead a subject to prefer any particular item of feedback. A number of different arrangements for the six feedback types were explored as shown in figure 32.

Figure 32: Different arrangements for the six feedback types



However all such arrangements were deemed to contain an inherent order, as shown in figure 33,

Figure 33: Inherent orders in the different arrangements



As such, the prime consideration was usability and therefore feedback items are presented to users in a vertical list as shown in figure 34.

Figure 34: A screenshot of the interface presented to subjects upon answering a question

3.6.2.6 Stage four: Implementation

Choice of technology

The technology chosen was Adobe Flex in conjunction with a Microsoft Enterprise Web Server and SQL server database. This choice of technology fulfilled the requirement that the testing instrument be Internet based and contains the flexibility required to build a bespoke tool. While similar technologies exist to Microsoft Enterprise Web Server and SQL server, these technologies were already in use at the host institution and were therefore the pragmatic choice.

Coding

As described above, my intention was to code the software tool myself however upon soon after starting this activity it became clear that this was not possible within the time allowed and my initial work was passed to the software developer. All code within the final deployed version was written by the software developer.

3.6.2.7 Stage five: Software testing

The focus of testing was prevent errors occurring prior to subjects using the tool rather than reacting to errors as subjects did use the tool. This was due to the importance of capturing accurate data during the periods that subjects were using the software tool. All testing was ‘dynamic’ in that it comprised testing working models of the software as opposed to static testing such as the examination of code. In accordance with the agile software development approach, testing was iterative and occurred throughout the development process. Each iteration comprised two stages:

Stage one: Testing of user cases

I conducted tests of pathways thought the system for each type of user.

Stage two: Testing with small groups of students

I asked groups of between two and three students to use the software and report on any errors they encountered or suggestions they had.

Following each stage I also conducted an analysis of the data that had been captured and conducted manual calculations of descriptive statistics produced by the tool in order to confirm their accuracy.

3.6.2.8 Stage six: Deployment

Once a testing stage was complete I signed off the system and then installed and activated the system.

3.6.2.9 Continued development: Versioning

As outlined earlier in this thesis under the section “Analysis of quantitative data: Data processing” (See page 93), the exploratory nature of this research required regular additions to the data set. These additions included new fields relating to the subjects such as ‘learner level’ and new items of TBI such as the “Accounting Primer” and “Pre-study Mathematics” courses.

While the question and feedback screens presented to subjects remained unchanged throughout, these additions did affect components of the software tool such as the registration screen, the administration screens, the database architecture, the algorithms and the internal logic of the system. As such, a new version of the software tool was developed for each study and development continued on the tool throughout the two and a half year period of data collection.

This concludes the description of the development of the software tool.

4. RESEARCH INTEGRITY

This chapter discusses measures taken to ensure the integrity of the research. Frameworks used to judge the integrity of research tend to be guided by principles. For example the host institution for the present research has adopted the Council for Science and Technology's 'universal ethical code for scientists' (CST 2006) which is guided by the following three principles :

- a. Rigour, Honesty and Integrity.
- b. Respect for Life, the Law and the Public Good.
- c. Responsible Communications: Listening and Informing

A different code, the Singapore Statement on Research Integrity (Steneck, Mayer et al. 2010) states four basic principles required for responsible research practice,

- a. Honesty in all aspects of research
- b. Accountability in the conduct of research
- c. Professional courtesy and fairness in working with others
- d. Good stewardship of research on behalf of others

(Steneck, Mayer et al. 2010)

The Singapore Statement also outlines 14 professional responsibilities held by researchers to ensure the integrity of their research. The following section uses these how each of these responsibilities were upheld over the duration of the present research.

1. *"Integrity: Researchers should take responsibility for the trustworthiness of their research." (Ibid. 2010)*

This thesis is an accurate record of the research.

2. *"Adherence to Regulations: Researchers should be aware of and adhere to regulations and policies related to research." (Ibid. 2010)*

The regulations of most note to the present research relate to the privacy of the research subjects. Privacy is protected under law by the European Convention on Human Rights and the European Data Protection Directive (95/46/EC). These are both reflected in the UK law in the Data Protection Act 1998 (Steneck, Anderson et al. 2012). This act of parliament states a list of legal obligations held by those who handle personal information.

Quantitative data are gathered in this research as subjects, who are students at the host institution, interact with elements of online courses accessed via the institution's virtual learning environment. These online courses form part of subjects' regular programme of study and are not specific to this research. All materials are delivered via the institution's IT infrastructure and all data are stored on the institution's databases. The College fulfils its obligations stated in Data Protection Act 1998 and the present research does not introduce any additional factors that jeopardize this compliance. Quantitative data were removed from the institution's IT infrastructure in order to be processed via statistical packages. At this stage data records were anonymised and these exported data are therefore not regulated by the Data Protection Act 1998.

Qualitative data were gathered in this research during the semi-structured interviews. These data were also stored anonymously and are therefore not regulated by the Data Protection Act 1998.

3. *"Research Methods: Researchers should employ appropriate research methods, base conclusions on critical analysis of the evidence and report findings and interpretations fully and objectively. " (Ibid. 2010)*

I have endeavoured to uphold this responsibility throughout the research process.

4. *"Research Records: Researchers should keep clear, accurate records of all research in ways that will allow verification and replication of their work by others. " (Ibid. 2010)*

A research record should contain the research data, information relating to how the data were collected and the thoughts and conclusions that occurred while conducting the research. (Steneck, Anderson et al. 2012).

In the present research, quantitative data were gathered periodically from the software system and subject to analysis in statistical software packages. On each occasion the source data were stored in a single folder with a descriptive title together with all files relating to the statistical software. This approach allows the statistical tests reported in this thesis to be readily repeated. Qualitative data were gathered during the semi-structured interviews. Audio and text transcripts of these interviews are stored in a single folder together with files relating to the CAQDAS software used for analysis.

A narrative relating to work being conducted, together with thoughts and conclusions, were documented in the format of a draft thesis from the outset of the research. A version of this thesis was saved following each significant addition or revision. Together, these theses form the majority of the research record.

5. *"Research Findings: Researchers should share data and findings openly and promptly, as soon as they have had an opportunity to establish priority and ownership claims. " (Ibid. 2010)*

A peer reviewed paper outlining the results of study one was presented at the 2011 IADIS Internet Technologies & Society. This paper won the award for Outstanding Paper. Following the submission of this thesis, the results of studies two, three and four will be prepared for submission to peer reviewed journals.

6. *“Authorship: Researchers should take responsibility for their contributions to all publications, funding applications, reports and other representations of their research. Lists of authors should include all those and only those who meet applicable authorship criteria. ” (Ibid. 2010)*

Steneck, Anderson et al. (2012) state an author to be somebody who has made a ‘substantial intellectual contribution’ to the research. My judgment is that I am the sole author of this research.

7. *“Publication Acknowledgement: Researchers should acknowledge in publications the names and roles of those who made significant contributions to the research, including writers, funders, sponsors, and others, but do not meet authorship criteria. ” (Ibid. 2010)*

Two people have made a significant contribution to this research. First, my supervisor Dr Benita Cox who provided strong support and guidance throughout. Second, Mr John Bicknell who assisted in the development of the software tool. Mr Bicknell’s contribution is detailed in in the section “The Development of the Software Tool” (See page 112).

8. *“Peer Review: Researchers should provide fair, prompt and rigorous evaluations and respect confidentiality when reviewing others' work. ” (Ibid. 2010)*

This responsibility is not relevant in the context of this thesis.

9. *“Conflict of Interest: Researchers should disclose financial and other conflicts of interest that could compromise the trustworthiness of their work in research proposals, publications and public communications as well as in all review activities.” (Ibid. 2010)*

I do not have financial or other conflicts of interest relating to this research.

10. *“Public Communication: Researchers should limit professional comments to their recognized expertise when engaged in public discussions about the application and importance of research findings and clearly distinguish professional comments from opinions based on personal views. ” (Ibid. 2010)*

This responsibility has not been relevant in the context of this thesis.

11. *“Reporting Irresponsible Research Practices: Researchers should report to the appropriate authorities any suspected research misconduct, including fabrication, falsification or plagiarism, and other irresponsible research practices that undermine the trustworthiness of research, such as carelessness, improperly listing authors, failing to report conflicting data, or the use of misleading analytical methods. ” (Ibid. 2010)*

I am not aware of any irresponsible research practice relating to this research.

12. *“Responding to Irresponsible Research Practices: Research institutions, as well as journals, professional organizations and agencies that have commitments to research, should have procedures for responding to allegations of misconduct and other irresponsible research practices and for protecting those who report such behavior in good faith. When misconduct or other irresponsible research practice is confirmed, appropriate actions should be taken promptly, including correcting the research record. ” (Ibid. 2010)*

This responsibility is not relevant in the context of this thesis.

13. *“Research Environments: Research institutions should create and sustain environments that encourage integrity through education, clear policies, and reasonable standards for advancement, while fostering work environments that support research integrity. ” (Ibid. 2010)*

This responsibility is not relevant in the context of this thesis.

14. *“Societal Considerations: Researchers and research institutions should recognize that they have an ethical obligation to weigh societal benefits against risks inherent in their work. ” (Ibid. 2010)*

Although this research received no direct public funding, a proportion of the income of the host institution derives from public funds. Steneck, Anderson et al. (2012) propose that public funds are provided with the expectation that the subsequent research will educate and benefit the public providing the funds. The present research fulfils this responsibility by revealing new insights into the educational process and contributes to the wider body of knowledge that helps educators design more effective learning environments.

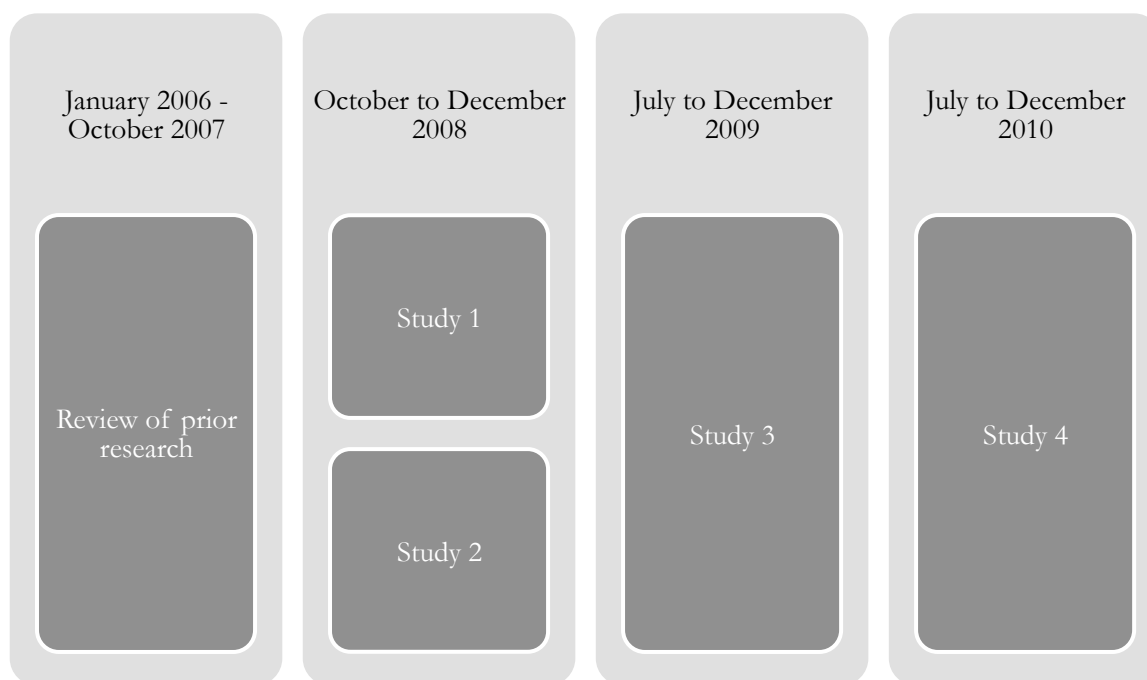
4. THE FOUR STUDIES: DETAILED DESIGN, RESULTS AND ANALYSIS

4.1 OVERVIEW OF THE FOUR STUDIES

This research adopted an exploratory design consisting of four studies over a period of three years involving 24 cohorts of students studying at one host institution. Studies one, three and four comprise three technology based observations. Study two comprises a series of semi-structured interviews.

In this research, the opportunity to collect data ran annually from August to October in each of the years 2007 to 2010. August to October being the period when the items of TBI are studied by students at the host institution. The four studies were conducted according to the schedule shown in figure 35.

Figure 35: The timeline of this research



Detailed reports on each of studies one, two, three and four now follow.

4.2 STUDY ONE: ESTABLISHING BASE DATA (2008)

4.2.1 Introduction

This study comprised a technology based observation of one cohort of students studying one item of TBI. The study took place in October 2008.

Subjects studied an item of TBI covering the topics 'Developing an Information Systems Strategy' and 'Managing Ethical Issues in Information Systems'. The TBI comprised presentational materials plus a series of multiple choice questions. The presentational materials were made available to the students online for a period of two weeks. At the end of this period, subjects took the multiple choice question component of the TBI together, in a computer laboratory and under supervision.

During the multiple choice question component subjects were presented with six feedback options after answering a question, these corresponding to the six types in the taxonomy of feedback. Subjects were also given the choice either to view feedback immediately following their response to a question or after having answered all of the questions. A software tool recorded subjects' interactions with this activity by recording the target and timing of each 'click' subjects made in the interface.

This study aims to test the hypotheses which emerged from the literature review regarding learners' preferences for TBI feedback. The hypotheses are,

1. Learners will elect to view feedback more frequently than they do not
2. Learners will elect to view KCR and KR type feedback more frequently than they elect to view either EX, IB or EI type feedback
3. Learners will elect to view immediate feedback more frequently than they elect to view delayed feedback
4. Learners' choice to view a feedback type is dependent on:
 - a. Whether a question is answered correctly or incorrectly (Factor #I51 Performance)
 - b. Gender (Factor #I6 Gender)

4.2.2 Formation of hypotheses

The development of hypotheses followed the following path. First, the literature review identified three general hypotheses regarding learners' preferences for the format of TBI feedback and these were chosen for investigation in study one.

1. Learners will elect to view feedback more frequently than they do not

2. Learners will elect to view KCR and KR type feedback more frequently than they elect to view either EX, IB or EI type feedback
3. Learners will elect to view immediate feedback more frequently than they elect to view delayed feedback

Then, potential factors were considered. The literature review identified nine factors, shown in table 8, that potentially affect whether a learner elects to view a feedback type. Seven of these were classified as being internal to the learner and two classified as being external, or environmental, variables.

Table 8: Factors that potentially affect whether a learner elects to view a feedback type

Internal factors	External factors
Factor #I1 Learner expectations	Factor #E1 Assessment
Factor #I2 Response certitude	Factor #E2 Feedback level
Factors #I3 Interpretation of task	
Factor #I4 Level of expertise	
Factor #I5: Performance	
Factor #I6 Gender	
Factor #I7 Error type	

Investigating all nine factors was beyond the scope of a single study. Two factors were chosen via the following process.

Factor #I7 error type and factor #E2 feedback level

First, factors ‘#I7 Error type’ and ‘#E2 Feedback level’ were eliminated as factors for investigation. Investigating error type was not possible due to the decision to conduct the study in an authentic educational context. Creating an accurate taxonomy of error type requires precise control of the learning environment. Taxonomies adopted in prior research (Birenbaum and Tatsuoka 1987; Dempsey and Litchfield 1993) were fine grained and applicable to only a narrow range of subjects and question types. In fact, the error type taxonomies required both Birenbaum and Tatsuoka and Dempsey and Litchfield to recreate the learning materials that formed the basis of their studies. Such an approach would not satisfy the requirement for an authentic educational context.

Factor ‘#E2 Feedback level’ was not incorporated for practical reasons. An analysis of the TBI forming the focus of study one revealed that all feedback was at the level of ‘task learning’ according to the categorization of feedback level proposed by Kluger and DeNisi (1996).

Factor #I1 Learner expectations, factor #I2 Response certitude and factor #I3 Task

Next, factors ‘#I1 Learner expectations’, ‘#I2 Response certitude’ and ‘#I3 task’ were removed from consideration as the technology based observation approach, adopted as the data collecting instrument in this study, is not an effective way to capture data relating to such internal factors. It may have been technically possible to incorporate

measurement of these factors into the user interface. However, even assuming this could be achieved, doing so would require additional meta-level questions to be inserted into the TBI which would compromise the authenticity of the materials.

Factor #E1 Assessment

The technology based instruction adopted in study one was not assessed, therefore factor ‘#E1 Assessment’ was eliminated as a focus of study.

Factor #I4 Level of expertise

This left three factors for consideration. ‘#I4 Level of expertise’, ‘#I6 Performance’ and ‘#I7 Gender’. Data relating to ‘#I4 Level of expertise’ were not available at the time of study one. This left two factors available for incorporation into hypothesis four, ‘#I6 Performance’ and ‘#I7 Gender’.

As a result, the hypotheses for study one were:

1. Learners will elect to view feedback more frequently than they do not
2. Learners will elect to view KCR and KR type feedback more frequently than they elect to view EX, IB or EI type feedback
3. Learners will elect to view immediate feedback more frequently than they elect to view delayed feedback
4. Learners’ choice to view a feedback type is dependent on:
 - a. Whether a question is answered correctly or incorrectly (Factor #I6 Performance) and
 - b. Gender (Factor #I7 Gender)

4.2.3 Design

The subject group comprised an existing class of 160 postgraduate students studying a 20-hour masters level course in Business Informatics as part of a postgraduate degree in general management. As described in the section ‘Research methods’ (See page 84), subjects throughout this research comprised authentic student cohorts at the host institution. This cohort of students was chosen as they were the only cohort who studied this particular Business Informatics course in 2008.

As described in the section ‘The three items of Technology Based Instruction (TBI)’ (See page 102), the TBI covers two topics from this course, ‘Developing an Information Systems Strategy’ and ‘Managing Ethical Issues in Information Systems’. The TBI comprised three hours of self-study and one of the 20 class-based taught hours. The TBI contained two components, online presentational materials in the form of scanned extracts from the

course textbook plus a multiple choice question based test. The latter test was delivered in class and subjects were asked to study the presentational materials in advance.

Subjects took the multiple choice test component individually using a personal computer in a computer laboratory. The 160 subjects were divided into two groups of 80 subjects and each group took the activity at a different time but on the same day. Data relating to the interactions of 146 subjects were obtained, the shortfall being primarily explained by a server side technical error resulting in approximately half the subjects from the first group being unable to take the activity. Normal classroom conditions, as opposed to examination type conditions, prevailed.

Subjects were told that their scores in the multiple choice questions would not contribute towards their course assessment and that the aim of the activity was to promote learning of the material. Care was taken to mitigate the presence of a time limit as a possible confounding factor. The activity is designed to take between 25 and 30 minutes to complete and subjects were given one hour to do so. In addition, subjects were told that they were not required to answer all of the questions. Four staff members, including myself and the course lecturer, were present throughout the session and subjects were encouraged to contact a staff member if they required assistance. Subjects were told they could leave the classroom when they had finished.

An unforeseen occurrence was that some subjects made multiple attempts at the exercise. Of these multiple attempts, a significant number appeared to be false starts in that students studied between one and three questions, stopped and then started the activity again. In order to remove bias in the data resulting from multiple submissions, including such false starts, only one submission from each subject was selected for analysis. This procedure is described in more detail in the section 'Analysis of quantitative data: Data processing' (See page 93).

The software tool developed as part of the research delivered the activity and tracked subjects' interactions with the instructional material. The software tool stored the time and target element of each 'click' made by a subject during the activity.

4.2.4 Selection of statistical tests

As described in the section 'Data analysis: Quantitative data' (See page 85), descriptive statistics and multilevel logistic regression were chosen as the statistical methods to determine whether the hypotheses were supported.

4.2.5 Reliability and Validity

General issues relating to the reliability and validity of the research design are introduced and discussed in the section 'Research methods' under 'Reliability and validity' (See page 96). As discussed in this section, the internal validity of the logistic regression analysis used to test hypothesis four is compromised by the possible presence of confounding factors.

Over the course of this research, 13 factors were identified which potentially affect learners' interactions with feedback. Two of these factors are included in the logistic regression analysis in study one: factors '#I5

Performance’ and ‘#I6 Gender’. Five factors remain constant and therefore do not need to be considered: factors ‘#E1 Assessment’, ‘#E2 Feedback level’, ‘#E3 Time limit’, ‘#E4 Location’ and ‘#E6 Subject matter’. One factor ‘#I4 Level of expertise’ was found not to be significant in study four.

However the five remaining factors potentially confound the results of the logistic regression analysis in this study. Three of these are affective factors internal to the subject: factors ‘#I1 Learner expectations’, ‘#I2 Response certitude’ and ‘#I3 Interpretation of task. A further factor ‘#E5 Question focus’ was unknown at the time of study one but is incorporated into subsequent analysis of the same data in study three. Factor #I7 Error type was not incorporated into the analysis. It was not possible to control for these factors in study one for the reasons explained above in the section ‘Formation of hypotheses’. The presence of these factors is noted in the later section ‘Critique of the design and conduct of the research’ (See page 268).

A further issue relating specifically to this study was that subjects comprised a single student cohort and focused on one subject area of their degree programme. As such the investigation cannot claim a high level of external validity. Caution should be applied when generalising results to other cohorts, particularly at different institutions and in different subject areas. Over the course of this research, this issue was mitigated to a degree by repetition. In subsequent studies, the observation was repeated with different cohorts of students studying different subjects.

4.2.6 Results

Investigation one was conducted in October 2008. Table 9 provides summary data relating to the study that will be referred to in subsequent sections.

Table 9: Summary data relating to study one

Number of students	146
Total number of questions answered	8016
Average number of questions answered per student	47.92
% of questions answered correctly	65.64%
% of questions answered incorrectly	34.36%

The results for each hypothesis were as follows.

4.2.6.1 Hypothesis 1: Learners will elect to view feedback more frequently than they do not

Table 10 below shows the mean, median and mode α figures for each feedback type together with the standard deviation. To illustrate, the table shows that on average students clicked on feedback type ‘KR for 84.47 per cent of the questions to which they responded.

The table also shows that subjects did not view feedback for 14.81 per cent of the questions they answered and therefore viewed feedback for 85.19 per cent of the questions they answered. Hypothesis 1 is supported.

Table 10: Summary statistics relating to α figures for each feedback type

	No feedback	KR	KCR	IB	EX	EI1	EI2
N	146	146	146	146	146	146	146
Mean	0.1481	0.8447	0.3223	0.0655	0.0889	0.0203	0.0191
Median	0.0000	1.0000	0.3200	0.0200	0.0200	0.0000	0.0000
Mode	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Std. Deviation	0.3259	0.3336	0.2124	0.1194	0.1743	0.0573	0.0642

While subjects did elect to view feedback more often than they did not, the figure of 14.81 per cent is of note.

4.2.6.2 Hypothesis 2: Learners will elect to view KCR and KR type feedback more frequently than they elect to view either EX, IB or EI type feedback.

As shown in table 10 above, the α value for IB type feedback is 0.0655 suggesting that, on average, subjects elected to view the relevant section of the original instruction for 6.55 per cent of the questions they answered. This α value for EX type feedback is 0.0889 and for EI1 and EI2 type feedback is 0.0203 and 0.0191 respectively. The α value for all four feedback types is lower than that of both KR and KCR type feedback and hypothesis 2 is supported.

The data also reveal further relationships between these feedback types.

IB and EX versus EI1 and EI2 feedback types

A distinction can be drawn between the IB and EX feedback types and the EI1 and EI2 feedback types. The α values for EI1 and EI2 type feedback are noticeably lower than that for IB and EX. The values of EI1 (0.0203) and EI2 (0.0191) mean that, on average, students looked at each feedback type just once.

Relationship between IB and EX type feedback

The data collected allow the relationship between IB and EX to be explored further. Both these feedback types can be considered to provide the information a subject requires in order to close any gap between what had been understood and what was required to be understood in order to answer the question correctly. Given this it is hypothesised that subjects will elect to view either one of these feedback types but not both. Table 11 below shows that this is generally the case. Between them, the students answered a total of 8016 questions and the column labelled 'frequency' shows the number of questions for which the IB and EX feedback types were selected. Subjects selected either IB or EX type feedback on 9.44 per cent of occasions but selected both IB and EX type

feedback on 1.97 per cent of occasions. Thus subjects were noticeably more likely to elect to view either IB or EX type feedback as opposed to both these feedback types. types.

Table 11: Relationship between IB and EX type feedback

	Frequency	%	Cumulative %
Viewed neither IB nor EX	7101	88.59	88.59
Viewed IB only	304	3.79	92.38
Viewed EX only	453	5.65	98.03
Viewed both IB and EX	158	1.97	100
	8016	100	

Relationships between KR and KCR type feedback

There is also an overlap in the information provided by the KR and KCR feedback types in that the former is a subset of the latter. KR type feedback tells the subject whether their response was correct or incorrect. KCR type feedback tells the subject the correct response. Thus KCR type feedback provides all the information contained in KR type feedback. It is therefore reasonable to consider that some subjects will elect to just KCR type feedback. However, table 12 shows that this is not the case. Subjects rarely elected to view just KCR type feedback.

Table 12: Relationship between KR and KCR type feedback

	Frequency	%	Cumulative %
Viewed neither KR nor KCR	1331	16.6	16.6
Viewed KR only	4306	53.72	70.32
Viewed KCR only	47	0.59	70.91
Viewed both KR and KCR	2332	29.09	100
	8016	100	

Of the 8,016 questions answered between the students, 2,754 were answered incorrectly. When just these questions are considered the relationship between KR and KCR becomes clearer. As shown in table 13, when subjects answer a question incorrectly then they generally elected to view both KR and KCR type feedback.

Table 13: Interplay between KR & KCR type feedback (Questions answered incorrectly)

	Frequency	%	Cumulative %
Viewed neither KR nor KCR	412	14.96	14.96
Viewed KR only	272	9.88	24.84
Viewed KCR only	29	1.05	25.89
Viewed both KR and KCR	2041	74.11	100
	2754	100	

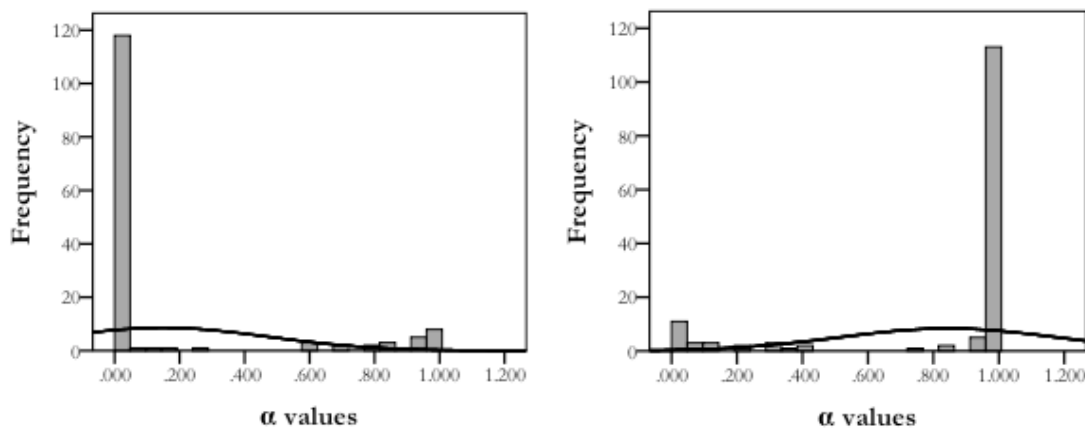
Analysis of the distribution of α values

Table 10 above presents summary statistics relating to the distribution of subjects' α values. When the distribution of these α values is considered, two points of interest emerge. These are now discussed.

Point one: Relatively high values of standard deviation for KR and KCR feedback types and the 'no feedback' condition

In table 10, the relatively high values of the standard deviation for the KR and KCR feedback types, and also the no feedback condition, are of note. Figure 36 below shows two histograms plotting the frequency distributions relating to KR type feedback and the 'no feedback' condition. The histograms plot the α value on the x-axis and the frequency of these α values on the y-axis.

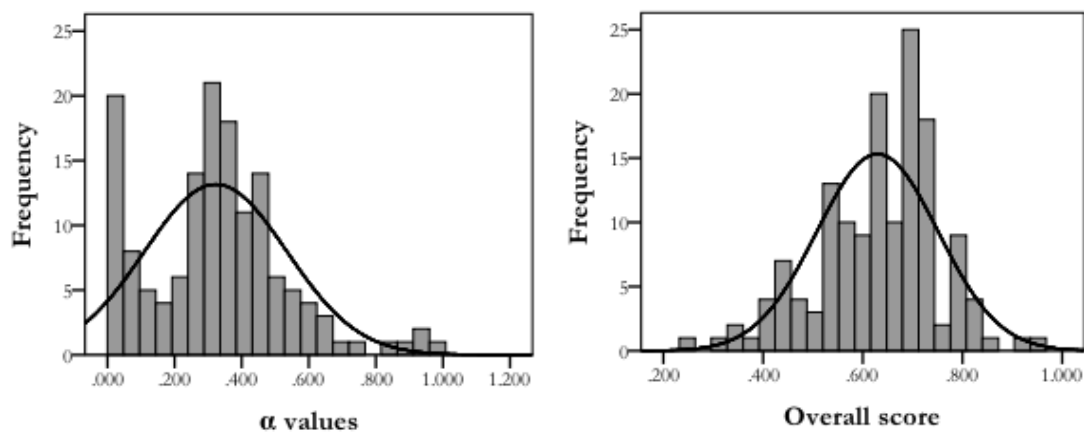
Figure 36: Distribution of α values for 'No feedback' (left) and KR type feedback (right)



These two histograms are, to a large degree, mirror images of each other. The single tall bar indicating that subjects did view KR type feedback for all of the questions reflects the tall bar indicating that subjects did not choose the 'no feedback' condition. In each case, the relatively large standard deviation is due to a cluster of students, at the opposite ends of the range to the mode. A minority of subjects exhibited notably different behaviour to the majority.

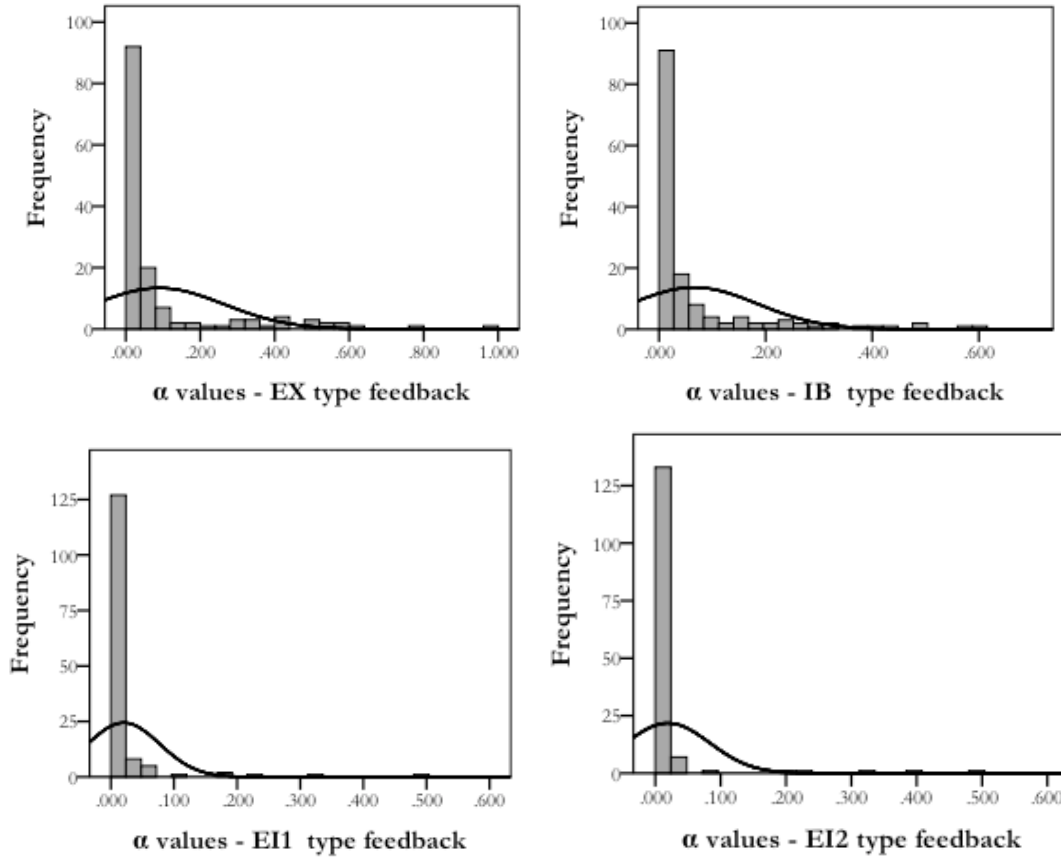
The equivalent histogram for the KCR feedback type is shown in figure 37 below together with a histogram showing the frequency distribution of student scores across the test. The positive skew in the former is largely mirrored by the negative skew in the latter. As will be demonstrated later, this is due to the fact that the selection of KCR is closely related to whether a question is answered correctly. The relatively large standard deviation in the α values for KCR type feedback is likely due to the spread of values across student scores.

Figure 37: Distribution of α values for KCR type feedback (left) and subjects' scores (right)



Point two: The distribution of α values for elaborative feedback types contains a marked positive skew. Subjects' α values cover a wide range for all elaborative feedback types however, as discussed in the section 'Research methods' (See page 87), the distributions of α values for these feedback types contain a strong positive skew. In table 10 this is shown by the median α value being significantly lower than the mean for these feedback types. A histogram showing the distribution of α values for KCR type feedback is shown above (figure 10), histograms for feedback types IB, EX, EI1 and EI2 type feedback are shown in figure 38 below.

Figure 38: Distribution of α values for EX/IB/EI1/EI2 type feedback



Unlike the histogram for KCR type feedback, the distributions of α values for feedback types IB, EX, EI1 and EI2 do not closely reflect the histogram plotting subjects' performance across the activity. This is because performance is less of a factor in determining whether subjects elected to view these feedback types and this is discussed under 'hypothesis four' below.

As mentioned above, it is also of note that a significant proportion of subjects did not view the elaborative feedback types at all. This is shown in the table 14 below.

Table 14: Subjects who did not view the elaborative feedback types

	KR	KCR	IB	EX	EI1	EI2
Never viewed	7	13	56	67	87	95
	4.79%	8.90%	38.36%	45.89%	59.59%	65.07%
Viewed	139	133	90	79	59	51
	95.21%	91.10%	61.64%	54.11%	40.41%	34.93%
Total students	146					

4.2.6.3 *Hypothesis 3: Learners will elect to view immediate feedback more frequently than they elect to view delayed feedback*

Table 15 below shows that subjects' preference for feedback timing changed during the exercise. Prior to starting the exercise, subjects were given the choice of viewing feedback after responding to a question (immediate feedback) or viewing feedback at the end of the exercise (delayed feedback). The table shows that at this stage 47.95 per cent of the cohort elected to view delayed feedback and 52.05 per cent elected to view immediate feedback.

However, after starting the exercise subjects were able to change this setting at any time. Table 15 shows that 63.70 per cent of subjects elected to change this setting and this had a significant effect on subjects' feedback settings at the end of the exercise. 93.84 per cent of subjects ended the exercise electing to view feedback after responding to a question and just 6.16 per cent of subjects electing to view feedback at the end of the exercise.

Table 15: Data relating to delayed and immediate feedback

	Delayed or Immediate	Frequency	%
Feedback position at start	D	70	47.95%
	I	76	52.05%
Changed mind	D	53	36.30%
	I	93	63.70%
Feedback position at end	D	9	6.16%
	I	137	93.84%
Total		146	

In conclusion, subjects demonstrated a clear preference for immediate rather than delayed feedback. Hypothesis 3 is supported.

4.2.6.4 *Hypothesis 4: Learners' choice to view a feedback type is dependent on:*

- a. *Whether a question is answered correctly or incorrectly (#I6 Performance) and*
- b. *Gender (#I7 Gender)*

As described earlier in the section 'Research methods' (See page 89), a multilevel logistic regression was used to determine whether this hypothesis was supported. First, variance inflation factors (VIFs) were calculated to identify the presence of collinearity between the independent predictor variables. Results revealed low levels of collinearity as shown in table 16.

Table 16: VIF Factors

Variable	VIF	1/VIF
Gender	1.00	0.9999
Performance	1.00	0.9999
Mean	1.00	

Next univariate and multivariate multilevel logistic regression analyses were performed for each feedback type. The results are shown in table 17 on the following page.

Table 17: Results of multilevel logistic regression analyses

Feedback type	Variable	Category	Univariable		Multivariable	
			Odds Ratio (95% CI)	P-value	Odds Ratio (95% CI)	P-value
NF	Performance	Incorrect	1		1	
		Correct	1.35 (0.99, 1.84)	0.06	1.35 (0.99, 1.84)	0.06
	Gender	F	1		1	
		M	1.21 (0.14, 10.13)	0.86	1.19 (0.14, 10.06)	0.87
KR	Performance	Incorrect	1		1	
		Correct	0.81 (0.60, 1.11)	0.60	0.81 (0.60, 1.11)	0.19
	Gender	F	1		1	
		M	1.03 (0.14, 7.63)	0.98	1.03 (0.14, 7.62)	0.98
KCR	Performance	Incorrect	1		1	
		Correct	0.002 (0.001, 0.003)	<0.001	0.002 (0.001, 0.003)	<0.001
	Gender	F	1		1	
		M	1.32 (0.84, 2.08)	0.26	1.96 (0.79, 4.82)	0.15
IB	Performance	Incorrect	1		1	
		Correct	0.38 (0.30, 0.48)	<0.001	0.38 (0.30, 0.48)	<0.001
	Gender	F	1		1	
		M	1.36 (0.69, 2.69)	0.38	1.37 (0.69, 2.72)	0.38
EX	Performance	Incorrect	1		1	
		Correct	0.22 (0.18, 0.28)	<0.001	0.22 (0.18, 0.28)	<0.001
	Gender	F	1		1	
		M	1.14 (0.47, 2.72)	0.77	1.14 (0.46, 2.81)	0.78
EI1	Performance	Incorrect	1		1	
		Correct	0.90 (0.61, 1.32)	0.59	0.90 (0.61, 1.33)	0.59
	Gender	F	1		1	
		M	1.02 (0.56, 1.85)	0.95	1.02 (0.56, 1.85)	0.95
EI2	Performance	Incorrect	1		1	
		Correct	1.31 (0.84, 2.04)	0.24	1.31 (0.84, 2.04)	0.24
	Gender	F	1		1	
		M	1.03 (0.51, 2.10)	0.93	1.03 (0.51, 2.11)	0.93

Factor ‘#I6 Performance’

Factor ‘#I6 Performance’ was not found to be a significant factor for feedback types NF, KR, EI1 and EI2.

However it was found to be a significant factor for feedback types KCR, IB and EX. The odds ratios were 0.002, 0.38 and 0.22 respectively meaning that, for each of these feedback types, subjects were significantly less likely to view them when a question had been answered correctly.

This effect is particularly strong for feedback type KCR. The odds ratios for IB and EX are less strong meaning that while subjects were more likely to view these feedback types following an incorrect response, this was not a necessary condition.

The effect of performance can be also be seen by comparing α values for questions answered correctly with α values for questions answered incorrectly. These values are shown in tables 18 and 19 below.

Table 18: α values for each feedback type – Incorrect responses only

	NF	KR	KCR	IB	EX	EI1	EI2
N	146	146	146	146	146	146	146
Mean	0.1407	0.8506	0.7589	0.0940	0.1307	0.0162	0.0117
Median	0.0000	1.0000	1.0000	0.0000	0.0000	0.0000	0.0000
Std. Deviation	0.3213	0.3328	0.3697	0.2011	0.2608	0.0411	0.0480

Table 19: α values for each feedback type – Correct responses only

	NF	KR	KCR	IB	EX	EI1	EI2
N	146	146	146	146	146	146	146
Mean	0.1513	0.8426	0.3223	0.0638	0.0739	0.0295	0.0290
Median	0.0000	1.0000	0.3200	0.0000	0.0000	0.0000	0.0000
Std. Deviation	0.3292	0.3349	0.2124	0.1685	0.1796	0.1225	0.1249

Factor ‘#I7 Gender’

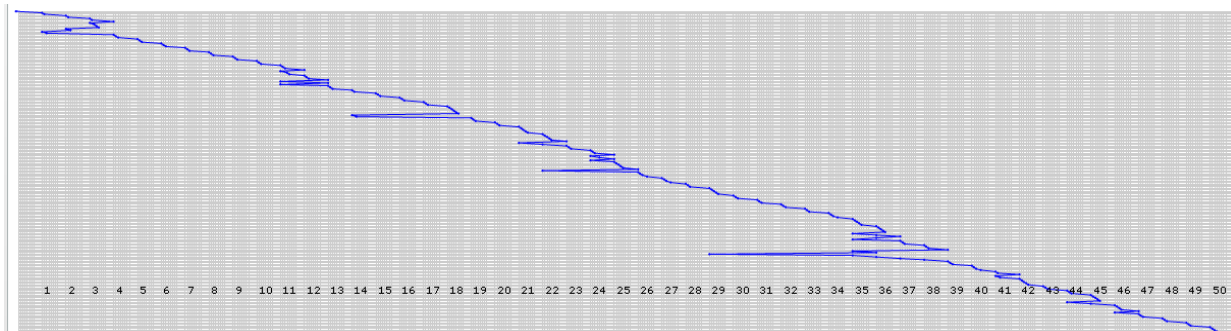
Factor ‘#I7 Gender’ was not found to be a significant factor for any of the feedback types. In addition to potentially being a factor that determined whether a subject chose to view a particular feedback type, gender was also potentially a factor in determining whether a subject elects to view immediate or delayed feedback. However, this appears unlikely from the data gathered in study one. Just seven subjects elected to view delayed feedback, three males and four females. While it is not possible to draw firm conclusions from such data, these data provide no evidence that gender is a factor influencing a preference for immediate or delayed feedback.

4.2.6.5 Further findings

Subjects revisited previous questions and viewed feedback 42.19 per cent of revisits.

The online system tracks subjects’ pathways through the test. Figure 39 below shows the pathway for one subject. If the subject had taken a straightforward linear pathway, moving from one question to the next in order, then we would see an approximately straight line from top left to bottom right. However, there are 15 horizontal diversions in figure 39, each of which indicates that the subject revisited a previous question.

Figure 39: An example of a subject pathway



An analysis of such diagrams for all subjects revealed that 86.30 per cent of students revisited at least one question over the course of the activity. Table 20 contains descriptive statistics for the cohort. On average, subjects revisited 4.17 questions. A majority of subjects revisited between one and four questions. The mean is larger than the median due to the 19 subjects (13.01%) who revisited nine questions or more and one subject in particular who revisited 29 questions.

Table 20: Number of questions revisited by subjects

n	146
Mean	4.17
Median	3.00
mode	2.00
Std. Deviation	4.17

The subjects were able to see previously viewed feedback types when revisiting questions without clicking on them. A more detailed analysis of the records for a random sample of 20 subjects shows that subjects elected to view an additional feedback item on 42.19 per cent of revisits.

Subjects' use of feedback did not change significantly over the course of the activity

Figures 40 to 44 below show the total number of clicks made by the subjects on each question first across all feedback types and then individually for feedback types KR, KCR, IB and EX. The 50 questions are listed in order along the horizontal axis. The vertical axis maps the total number of clicks. If subjects' use of feedback changed significantly over the course of the activity then we would expect to see an upward or downward trend in these figures however the variation remains consistent over the course of the activity. From this, it can be concluded that subjects' preference for feedback remained constant over the course of the activity.

Figure 40: Total clicks across all feedback types per question

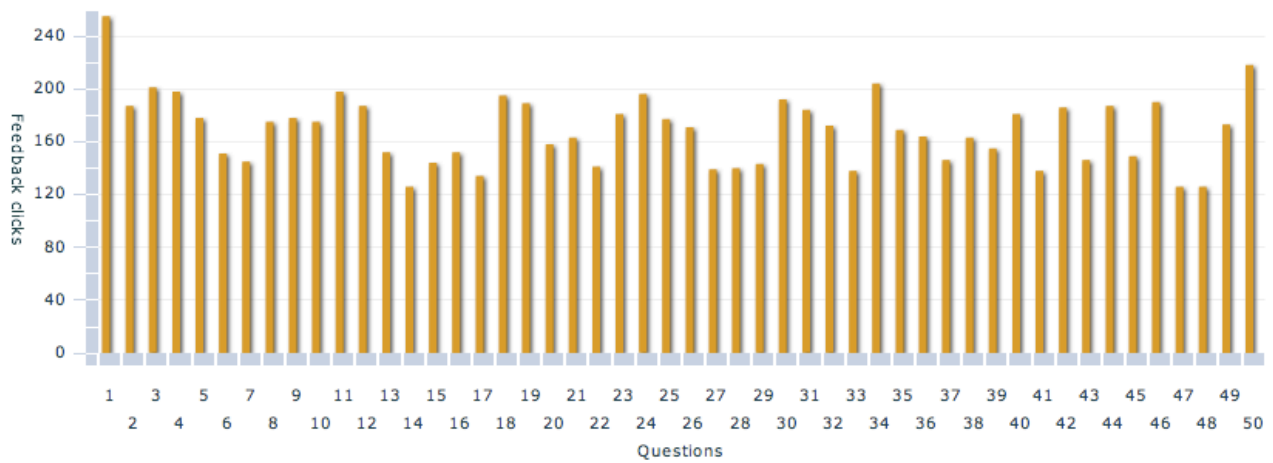


Figure 41: Feedback clicks for KR type feedback

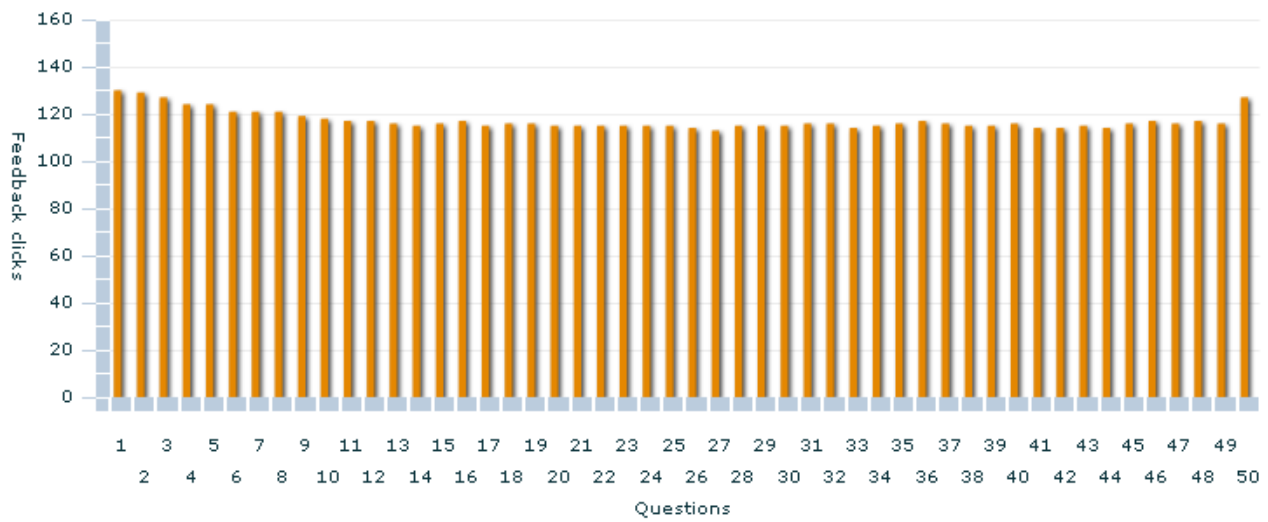


Figure 42: Feedback clicks for KCR type feedback

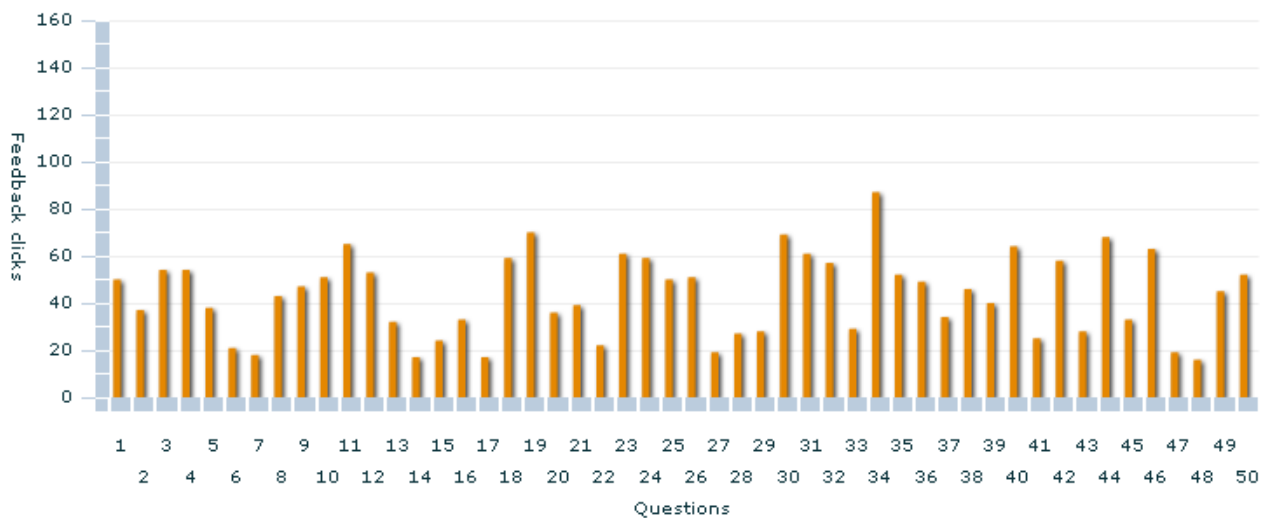


Figure 43: Feedback clicks for EX type feedback

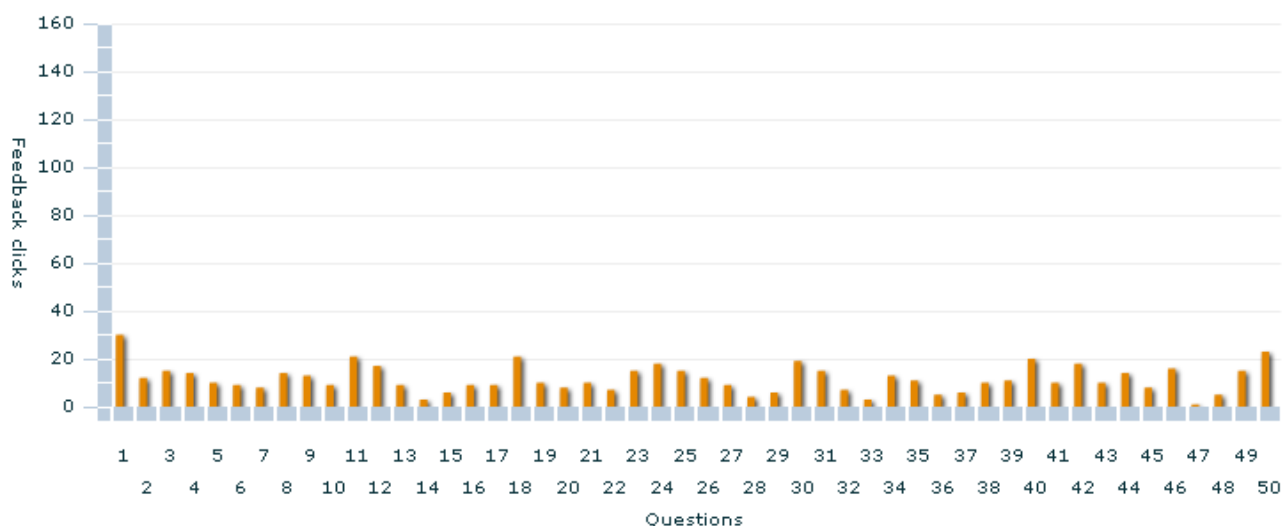
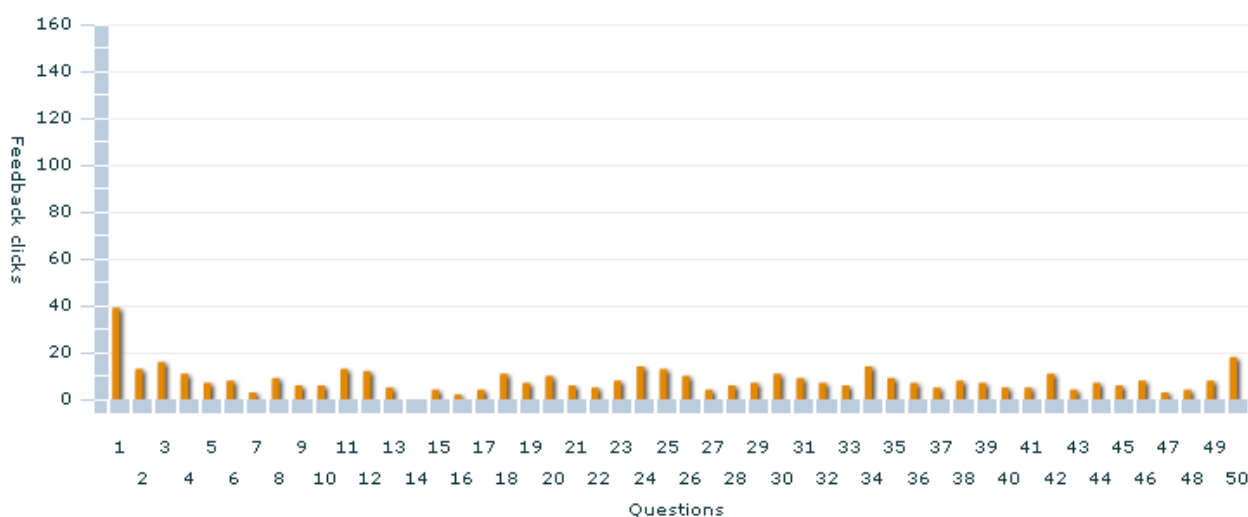


Figure 44: Feedback clicks for IB type feedback



4.2.7 Analysis and conclusions

This section organizes the results of study one into themes and relates these themes to the conclusions of previous studies introduced in the literature review.

4.2.7.1 *Subjects viewed feedback extensively.*

First, it is apparent that subjects viewed feedback extensively. Given the choice, subjects in this study did elect to view feedback. Subjects elected to view at least one feedback type for 85.19 per cent of the questions they answered. That said, it is notable that subjects did not elect to view any feedback for 14.81 per cent of questions, a not insubstantial proportion. In addition, the use of the six TBI feedback types varied considerably. Subjects elected to view KR and KCR type feedback more often than the other feedback types. This is discussed further in the following sections.

4.2.7.2 *The existence of a threshold*

Subjects elected to view feedback types KR and KCR more frequently than types IB and EX which in turn were viewed more frequently than types EI1 and EI2.

Prior studies have proposed the existence of a threshold beyond which the study of feedback becomes either ineffective or inefficient (Phye 1979; Kulhavy, White et al. 1985; Dempsey and Litchfield 1993). These three studies suggest that KCR type elaborative feedback represents this threshold and that viewing the other types of elaborative feedback is not perceived worthwhile by a learner. This result is consistent with that of Pridemore and Klein (1995) who found that subjects who received KCR plus EX type feedback were less likely to have enjoyed a lesson and less likely to feel they had 'learned a lot' than those receiving KCR type feedback alone.

While the findings of this study generally support the existence of a threshold at the level of KCR type feedback, this study also found that IB was viewed on average for 6.55 per cent of questions, EX for 8.89 per cent, EX for 2.03 per cent and EI2 for 1.91 per cent. The figures for EI1 and EI2 mean that, on average, subjects looked at these types for just one question. However, on average, subjects elected to view types IB and EX repeatedly meaning that were occasions in study one when this threshold was not present. The nature of these occasions is explored further over the course of this research.

The existence of a threshold is of interest to educational researchers, publishers of educational materials and teaching practitioners. Educational researchers have conducted studies investigating the effects of elaborative feedback on performance (See appendix four to this thesis), publishers have invested heavily in electronic learning systems that contain elaborative feedback systems based on electronic versions of their textbooks (Wiley 2009) and practitioners are encouraged to provide feedback to learners through, for example, the inclusion of elaborative feedback options in commonly used testing software.

Generally, elaborative feedback is more complex than verificational feedback and requires a significantly higher investment of time and resources to develop. Such feedback also requires a significantly higher investment of time and effort by the learner. An important question for educators is whether any resulting learning benefits justify this investment. Kulhavy et al. (1985) explicitly address the issue of feedback efficiency and conclude that elaborative feedback beyond KCR type feedback is not efficient and does not represent a worthwhile investment of learners' time.

Dempsey and Litchfield (1993) also attempt to determine the relative efficiencies of different feedback types adopting the same measure of efficiency developed by Kulhavy et al. (1985: 285).

$$\text{Feedback efficiency} = \frac{\text{correct retention percentage score}}{\text{square root of feedback study time}}$$

As per Kulhavy et al. this study suggests that students spent more time studying the elaborate types of feedback but also found no difference in performance in the retention test across the four conditions. Using the measure of

feedback efficiency above, they concluded that KCR feedback is the most efficient type of feedback among the four.

The data gathered in study one indicate that subjects perceived such a threshold and only elected to view elaborative feedback types beyond KCR on a minority of occasions.

4.2.7.3 Subjects held a preference for elaborative feedback beyond KCR that focuses specifically on the knowledge gap

One hypothesis that can be drawn from looking at the data across the feedback types is that subjects prioritise feedback that focuses on the specific gap between what was understood and what was required to be understood. This is termed the ‘knowledge gap’ in this research.

When considering the elaborative feedback types beyond KCR, EX type feedback was viewed most frequency (mean α value = 0.0889) followed by IB type feedback (0.0655) and then the two types of extra-instructional feedback, EI1 (0.0203) and EI2 (0.0191). It is proposed here that this order reflects the extent to which each of these feedback types focuses on the specific gap between what was understood and what was required to be understood.

- EX type feedback provides an explanation as to why the subject’s response was incorrect. This feedback type is most focused on the knowledge gap.
- IB type feedback presents a subject with the relevant section from the course material. This feedback type is sufficient to close the knowledge gap when a question recycles concepts verbatim from the course material. However, this feedback type may not be sufficient to close the knowledge gap when a subject is required to apply concepts from the materials to the question. In general, this feedback type is considered here to be less focused on the knowledge gap than EX type feedback.
- Extra-instructional or EI type feedback provide additional information relevant to a question.
 - EI1 type feedback provides further examples of the concepts introduced in a question or solutions to similar exercises. This information assists subjects in closing the knowledge gap but is less specifically focused than either IB or EX type feedback.
 - EI2 type feedback provides explanations of terminology and rare words. This is not generally sufficient to close the knowledge gap and is generally the least focused feedback type.

Thus the information provided in EI type feedback assists a learner to resolving a knowledge gap but, in itself, is not sufficient to close such gaps. This feedback type is thus less focused on a feedback gap than types EX and IB.

This would explain subjects' relatively low usage of the extra-instructional feedback types, EI1 and EI2. 119 of the 146 students did not elect to view EI1 type feedback during the activity. 127 students did not view EI2 type feedback. The mean α figures for these feedback types were lower than for types KR and KCR and also types IB and EX.

This distinction can also explain the varying impact of 'performance' across the elaborative feedback types. Whether a subject answers a question correctly or incorrectly, strongly affects their choice to view feedback types KCR, IB and EX but not EI1 and EI2.

4.2.7.4 The effect of Factor #15 Performance

One hypothesis drawn from the literature review was that learners would not choose to view TBI feedback following correct responses. The reason being that even if learners did elect to choose TBI feedback following a correct response then this action would be unlikely to improve their performance in subsequent tests.

While performance was found to be a significant factor in determining whether subjects elect to view a feedback type, the result was not absolute. Subjects were more likely to view feedback types KCR, IB and EX following an incorrect response but they also elected to view these feedback types following a correct response.

One explanation may lie in the response certitude theory proposed by Kulhavy and Stock (1989). For example, if a subject guesses the correct response then they may still be interested to view the information provided by explanatory feedback.

4.2.7.5 Patterns in subjects' behaviour

Consideration of subjects' behaviour in this study revealed two patterns. First, a majority of subjects elected to view feedback type KR for virtually all of the questions.

Second, the distribution of α values across the cohort contained a marked positive skew for each of the elaborative feedback types. A majority of subjects viewed these feedback types rarely and a minority viewed these feedback types at a wider range of frequencies.

For KCR type feedback this difference in behaviour was largely explained by the differences in performance across the cohort. However, for feedback types IB, EX, EI1 and EI2, factor '#15 Performance' is not able to totally explain these distributions. This will be investigated further in studies two, three and four.

4.2.7.6 Subjects prefer immediate feedback

The literature review revealed an on-going and inconclusive debate as to whether feedback should be provided immediately after a response or following a delay. Kulhavy and Stock (1989: 282) point to contemporary research indicating that a delay between the original instruction and the subsequent feedback leads to significantly better performance. Roper (1977) suggests that the effect may be more significant for some subjects. Roper found that

delayed feedback is more effective for learners with prior knowledge of a subject than for those studying for the first time. However, a number of researchers reject the existence of such an effect (Kulik and Kulik 1988; Dempsey, Driscoll et al. 1993; Dihoff, Brosvic et al. 2003).

In this study subjects expressed a clear preference for immediate feedback. 137 of the 146 subjects elected to view feedback immediately following their response to a question. This finding is consistent with that of Dempsey and Litchfield (1993) who found that,

“A second question asked learners to choose which statement concerning immediate feedback was most true for them. The largest percentage of learners (46%) responded that it satisfied their curiosity about whether the answer was correct. Another 33 per cent of the learners felt that it helped them learn information more quickly. Only 13 per cent replied that it encouraged them when they got the correct answer. Even fewer learners (7 per cent) answered that feedback was helpful, but could wait until the end of the lesson.” (Dempsey Litchfield 1993: 11)

Reasons for this preference will be investigated in study two.

4.2.7.7 Subjects revisit questions and feedback

86.30 per cent of students revisited at least one question over the course of the activity. This is of interest to those in the instructional design community who are required to make decisions regarding users' pathways through instructional material. If users prefer to return to previous screens then instructional designers may want to incorporate this functionality into the design of instructional materials.

4.2.8 Summary

The primary aims of study one were first, to test the hypotheses resulting from the literature review with an initial cohort of students and second, to identify results and hypotheses that could be tested in subsequent studies.

Hypotheses one to three were supported. The test of hypothesis one found that subjects viewed feedback extensively. However it was notable that subjects did not view feedback for 14.81 per cent of the questions to which they responded. The test of hypothesis two suggests the existence of a feedback threshold, at the level of KCR type feedback, beyond which subjects elected to view the elaborative feedback types noticeably less frequently. The test of hypothesis two also suggests subjects held a preference for elaborative feedback beyond KCR that focuses specifically on the knowledge gap. The result of the test of hypothesis three was unequivocal, subjects in this study demonstrated a strong preference for immediate feedback.

The test of hypothesis four found factor ‘#I6 Performance’, whether a subject responds correctly to a question, to be a significant factor affecting whether subjects chose to view feedback types KCR, IB and EX. However factor ‘#I7 Gender’ was not found to be a factor for any of the feedback types.

This study also revealed a number of patterns in behaviour from which hypotheses will be formed for the subsequent studies in this research. For example, the majority of subjects in this study chose to revisit at least one question during the activity.

These results will be investigated further in study two which aims to identify the reasoning behind the choices made by subjects in this study and the factors that affect these choices.

4.3 STUDY TWO: QUALITATIVE INTERVIEWS (2008)

4.3.1 Introduction

This study comprises a series of 10 semi-structured interviews and took place in November 2008. The overarching aim was to investigate why subjects made their choices in study one and the factors that affect these choices.

Specifically, the study investigated the following two questions,

1. Why do learners elect to view TBI feedback?
2. What factors affect whether learners elect to view a TBI feedback type?

The study comprised semi-structured interviews of 10 subjects selected from study one by random sample.

4.3.2 Design

4.3.2.1 Procedure

In the original design, the 10 interviews were to be held in the same week as study one. However, practical issues relating to subjects schedules and study load resulted in these interviews taking place approximately one month after study one. This issue is discussed further in the later section “Critique of the design and conduct of the research” (See page 268).

The interviews were conducted over two days, individually, in a closed room and with the activity available to both interviewer and interviewee for reference. Subjects were told in advance that they would be asked questions relating to their study of the TBI activity in study one. Due to the one month delay, subjects were instructed to refamiliarise themselves with the activity prior to the interview. Subjects were given no further instructions, nor were details relating to the content of the interview revealed. To avoid note-taking errors, responses were recorded and subsequently transcribed.

4.3.2.2 Selection of interviewees

All interview subjects were participants in study one. Practical constraints prevented all participants from study one being selected for interview. These included the time availability of the subjects and their willingness to participate. As such a sample was taken.

A key decision was the number of subjects to select for interview. Morse (2000: 3) proposes a number of factors that should be taken into account when estimating the size of a sample in qualitative research. These include:

“... the quality of data, the scope of the study, the nature of the topic, the amount of useful information obtained from each participant, the number of interviews per participant, the use of shadowed data, and the qualitative method and study design used.” (Morse 2000: 3)

An initial sample size of 10 was chosen. In order for the sample to be as representative as possible, in other words that the sample accurately reflected the larger population, potential subjects were initially selected via a simple random sample. 10 subjects were selected for interview at random from the list of all subjects who participated in study one. 'Random.org', an Internet based service which generates random numbers based on atmospheric noise, was used to determine the selected subjects.

However, only 6 of the 10 subjects initially selected agreed to be interviewed. A further 10 potential subjects were then selected from the remaining subjects on the list, via the same procedure, and invited for interview. Of these 10 subjects, 3 agreed to be interviewed. Finally a further 10 subjects were selected and the first subject to agree to be interviewed was chosen to fill the final place. As a result of this procedure the 10 subjects were, to a degree, self-selecting and study two is thus exposed to the risk of selection bias. For example, one potential bias is that the sample group of 10 subjects may have a stronger sense of motivation, on average, than the larger group of subjects who participated in study one. This issue is acknowledged in the section "Critique of the design and conduct of the research" later in this thesis (See page 268).

A further question is whether selecting 10 subjects was a sufficiently large number to reach robust conclusions. The key consideration was 'data saturation'. This term was originally associated with one particular type of sampling, 'theoretical sampling', and refers to the point at which no new insights are obtained should the sample size be increased (Ritchie and Lewis 2003: 81). Guest, Bunce et al. (2006: 78) conducted an analysis aiming to identify the number of interviews required for data saturation to be reached. They suggest 12 interviews can represent the data saturation point given certain conditions,

"For most research enterprises, however, in which the aim is to understand common perceptions and experiences among a group of relatively homogeneous individuals, twelve interviews should suffice."
(Guest, Bunce et al. 2006: 79)

In the same paper, Guest, Bunce et al. suggest 6 interviews provide sufficient data to support major themes. In the present research, all major themes had been identified following the analysis of 6 interviews and the data saturation point was considered to have been reached at 10 interviews. As such it was not necessary to select further subjects for interview.

4.3.2.3 The design of the topic guide

The topic guide used in the interviews is presented in appendix one of this document. The topic guide aims to elicit subjects' reasoning for electing to view TBI feedback and also to identify the internal and external factors that affect whether subjects' choose to view each TBI feedback type.

The topic guide contains a combination of open and closed questions. The open question format was chosen in order to prevent issues relating to prompting. The topic guide was piloted with three respondents, not from study one, in advance of the investigation.

When constructing the topic guide, a key decision was whether to include explicit references to the seven internal factors and two external factors, potentially affecting whether subjects chose to view a feedback type, that had been identified during the literature review (See page 59). Including explicit references would ensure that subjects' responses did include consideration of these factors. However, it was decided that doing so would potentially guide subjects to focus overly on these factors and that the inclusion of such questions would influence the outcome of the interviews to an unacceptable degree. Explicit references to these nine factors were not included in the topic guide.

The interviews were semi-structured and respondents were encouraged to explore tangential comments and discuss issues around the topics. In this way a number of themes became apparent during the series of interviews and these were explored further in subsequent interviews. In addition, a degree of exploration of answers specific to the questions was required in order to avoid superficial answers.

4.3.2.4 Method

As described in the earlier section 'Research methods', the explanation building method described by Yin (2002: 143) was adopted for the analysis of the interview transcripts. This method proceeds as follows:

1. Make an initial theoretical statement or an initial proposition
2. Compare this against the findings of an initial case
3. Revise the statement or proposition to match the initial case if required
4. Compare the revised statement or proposition against remaining details in initial case
5. Compare the revised statement against the findings of the remaining cases
6. Repeat this process as required

For each of the two questions an initial theoretical statement was determined and this was then gradually revised over the course of the analysis to match findings in the interview transcripts. The CAQDAS software, NVivo, was used to assist in this process. Prior to the analysis the nine potential factors identified in the literature review were coded as NVivo nodes.

4.3.2.5 Reliability and Validity

General issues relating to the reliability and validity of the research design are introduced and discussed in the section 'Research methods' under 'Reliability and validity' (See page 96).

The principal concern relating to reliability and validity in this study is that of internal validity. In semi-structured interviews, the primary threat to internal validity is that the questions in the testing instrument do not elicit the information intended. Approaches to mitigating this threat can be considered in relation to the three participating agents; the testing instrument, the research subjects and the interviewer.

First a great deal of consideration was given to the design of the testing instrument. Care was taken to ensure that questions were clear, focused and matched the stated aims of the investigation. In addition, as described above in section ‘The design of the topic guide’, steps were taken to ensure that questions were not phrased in a manner that guided subjects towards a particular response.

Second, potential causes of bias in subjects’ responses were considered and mitigated. Aside from the potential interviewer effects described shortly, the main identified potential cause of bias was any significant change in the background situation during the series of interviews. For example, if the subjects had sat an exam during the interview period then this may have affected the responses of subjects attending subsequent interviews. To mitigate this issue the interviews were conducted across two consecutive days, shortly after subjects completed their course and prior to their exam. As noted under ‘Procedure’ above, it was not possible to conduct the interviews until one month following study one due to subjects’ schedules and this resulted in a dilemma. If subjects were asked to repeat the activity then their responses would likely refer to their experience of repeating the activity and not their initial experience in study one. If subjects did not look at the activity prior to being interviewed then their responses would be compromised due to a lack of recall. As a compromise, subjects were told that they would be interviewed on their experiences in study one and were asked to familiarise themselves with the activity prior to being interviewed.

A third potential threat to the internal validity of the interviews was that of interviewer attribute effects. Interviewer attribute effects occur when the personal attributes of the interviewer affect interviewee responses and introduce bias into their responses. For example, the nature of the relationship between interviewer and interviewee may introduce such bias. In studies such as this, interviewees may tend towards positive responses if interviewed by the person that is known to have built the instructional activity. As such the use of an external objective interviewer was considered, however it was decided it would be valuable to conduct the interviews myself on this occasion in order to develop my ability in this area. To mitigate interviewer attribute effects, subjects were told that the interviews were anonymous and were encouraged to be forthright.

4.3.3 Results and analysis: Question one: What are the reasons learners elect to view TBI feedback?

The analysis of question one started with the initial theoretical statement that subjects chose to view feedback in order to achieve closure on the task. In Handley’s (2003) thesis, the concept of closure is considered the overarching driver behind subjects’ decisions to view feedback. Over the course of this analysis this statement was refined. When viewing different feedback types, subjects pursue closure first by seeking affirmation and second, by resolving knowledge gaps. One subject articulated this succinctly,

“Sometimes you want to be proved right, sometimes you want to understand...” [Subject #8]

These two steps are now described.

Pursuing closure: Step one - Subjects seek affirmation:

Subjects first sought affirmation. They sought confirmation that their response had been correct. The most common verb used by subjects to describe their need for affirmation was ‘want’. Subjects wanted to know whether they had responded correctly.

“Because I want to know whether the answer’s correct or not.” [Subject #5]

“I didn’t check everything actually, or mostly... and of course you want to know the answers, that was very...” [Subject #2]

“I really wanted to know whether I got those right or wrong.” [Subject #7]

As the following dialogue demonstrates, subjects found it difficult to articulate this want more precisely.

Respondent: “We wanted to know if we... which one was right because...” [Subject #1]

Interviewer: “Why?”

Respondent: “Because you wanted to know, you know, you wanted to know.” [Subject #1]

Interviewer: “Why?”

Respondent: “Just, rather it’s just so you know what’s right and what’s wrong, what you got wrong.”
[Subject #1]

Pursuing closure: Step two – Subjects aim to resolve knowledge gaps

Beyond affirmation, subjects generally chose to view feedback in order to resolve ‘knowledge gaps’ that had been identified while viewing the question and KR type feedback. A ‘Knowledge gap’ is defined here as any gap between what the subject understood and what had needed to be understood in order to have responded to the question correctly.

Subjects used stronger language to describe this requirement to resolve knowledge gaps. The word ‘need’ was typical.

“I would still look at the explanation if I got it wrong. And ... probably the book reference if I needed to.” [Subject #9]

“(If) it was like a very in-depth topic that, you know, you need to refer to a few other sources.”
[Subject #4]

“I kind of realised that... right or wrong and this could tell me where it’s at in the book. To me... that was all I, all I needed.” [Subject #10]

Generally, the requirement was to understand.

“Even though you got the wrong answer if you go to here then you understand.” [Subject #2]

(Regarding EI2) “It’s interesting but I don’t know if I’d really use it unless I just really didn’t understand something.” [Subject #10]

“So when I do it wrong, then I go to... see the right answer and I’d be like, okay I understand, then I wouldn’t do book reference. But I’d be like no, I don’t (understand) then I would have to go to book reference to prove it.” [Subject #8]

“And most of the time that was enough, but there was sometimes when you weren’t sure... you still weren’t sure why you it is incorrect.” [Subject #7]

“So most of the time you... if you’re wrong you would click on that, but then that’s enough, you kind of understand.” [Subject #9]

To a lesser extent, subjects stated a requirement to learn.

“If I happen to answer incorrectly, I want to gain the knowledge. That means I don’t know that answer so I want to know... I want to learn.” [Subject #6]

“I would just see further explanation, so say, to see if I can find more information, learn something new.” [Subject #8]

“(Looking at KCR type feedback) you go okay, right, so that’s, yeah, that’s the right answer you’d learn.” [Subject #1]

Two subjects referred to a connection between gaps in knowledge at the level of the TBI activity and gaps in knowledge at the wider course level. Their wish to close a gap resulted from a wish to increase their performance on the course as a whole.

“Yeah, I suppose you wanted to know for your studies. It’s useful to know the right answer and you’d use it for, you know, revision and stuff like that.” [Subject #1]

“I look for the fact that if I didn’t know an answer, why I didn’t know an answer. Where was it in the course material that I didn’t read about it? The other thing was that, or is a factor, if I happen to answer incorrectly; I want to gain the knowledge. That means I don’t know that answer so I want to know... I want to learn.” [Subject #6]

When articulating the achievement of closure, subjects used terms such as ‘being happy’, to ‘get it’, ‘to understand’, that the feedback ‘made sense’ or that their curiosity had been satisfied.

“After further thinking about it, seeing the correct answer, maybe I came to the conclusion that yes, okay, I get it.” [Subject #7]

“So if I get the right answer then I know what it’s talking about and didn’t necessarily check everything cos I was happy.” [Subject #2]

“I suppose that’s where my curiosity stopped.” [Subject #5]

Thus subjects chose to view feedback to seek affirmation and resolve knowledge gaps.

4.3.4 Results and analysis: Question two: What factors affect whether learners elect to view the TBI feedback types?

The literature review identified nine factors, shown in table 21, that potentially affect whether subjects choose to view a feedback type. These are categorized as being either internal or external to the learner.

Table 21: The nine potential factors identified in the literature review

Internal factors	External factors
Factor #I1 Learner expectations	Factor #E1 Assessment
Factor #I2 Response certitude	Factor #E2 Feedback level
Factors #I3 Interpretation of task	
Factor #I4 Level of expertise	
Factor #I5: Performance	
Factor #I6 Gender	
Factor #I7 Error type	

The analysis of transcripts to address question two started with the initial theoretical statement that these nine factors did not affect whether subjects chose to view a feedback type. For some factors this theoretical statement remained unchallenged, for others the statement was challenged and for the remainder no evidence was provided either way. In addition, the analysis revealed four additional external factors that subjects considered to affect whether they chose to view the feedback types. The findings for each of these 13 factors are now introduced and discussed.

4.3.4.1 Factor #I1 Learner expectations

The ten interviews provided strong support for the hypothesis that learners’ expectations are a factor affecting whether learners elect to view a feedback type.

That learners' expectations regarding feedback potentially determine whether a learner will elect to view a feedback type derives from the work of Handley (2003). Handley found that learners' achievement of closure is attained by meeting their expectations of feedback and that these expectations arise from four factors:

1. Learners' 'style of Engagement' with the instructional activity
Handley defines seven distinct styles ranging from 'mindful learning' to 'going through the motions'.
2. Learners' 'sense of inquiry' arising from the task
This refers to the learners' level of interest in the activity.
3. Learners' 'assessment of personal risk' arising from the task
This refers to learners' assessment of their ability to deal with the real life scenarios presented in Handley's study together with the implications of failure.
4. Learners' generic expectations of feedback
Broadly speaking, learners generally expect feedback to verify, correct, and explain.

In this series of interviews, subjects did not refer to an assessment of personal risk. This is to be expected as scenarios were not a significant part of the TBI activity, unlike the activities in Handley's study. Nor did subjects refer to any generic expectations of feedback. However, subjects did indicate that both their style of engagement and their sense of enquiry were factors determining whether they elected to view the feedback types.

Style of engagement

Handley (2003) proposes seven different styles of engagement that learners adopt while studying TBI. These styles are shown in table 22 below.

Table 22: Summary of Handley's (2003: 212) styles of engagement

Style-of-engagement	Description
Mindful learning	An intentional, thoughtful approach towards the e-learning material, and a desire to learn, rather than merely 'pass a test'.
Looking for take-aways	Learners adopting this approach were keen to learn quickly and efficiently, by accumulating behavioural rules, best practice guidelines, process checklists and other tools and take-aways for future reference.
Exploring	A playful approach to learning. When completing a task, learners sometimes selected a response option which they knew to be incorrect, out of a playful interest to see what the feedback might be.
Watching TV	Learners were emotionally immersed in the scenario, as though watching a TV drama unfold. Learners were passive in the sense that they were not mindfully trying to identify the key learning points.
Playing the game	Learners focused on working out what the application wanted from them, and then following that lead. By thinking in tune, they hoped always to select the correct response (whatever their own preferred choice) and finish the e-learning quickly and efficiently. This approach was actively

	instrumental.
Getting through it	An instrumental approach, whereby the learner wanted to 'get through' the e-learning application as quickly as reasonably possible.
Going through the motions	Learners made swift guesses at the right answer, but were not particularly worried whether their responses to activities were incorrect or correct. This was a passively instrumental style, even more so than 'getting through it'.

In this series of interviews, the subjects stated they had adopted four of the different of styles of engagement from Handley's taxonomy when taking the activity. In addition, subjects stated their style of engagement to be a determining factor when electing to view a feedback type. The four different styles were,

One: Playing the game

"I found (this type of learning) very efficient ... because it's like gaming, you know, you are so much involved.... it's fun first of all. It's like computer game and I found myself much more involved in this activity." [Subject #2]

Two: Going through the motions

Interviewer: "It sounds like your focus is not to learn it's to get to the end."

Respondent: "That's right." [Subject #5]

Respondent: "... you just want to have it done as quick as possible. It's the speed really so say if you want, you know, you want to have the answer straightaway and just the answer perhaps, if it's yes or no, or if it's one of the options you're given. Why did I move on? I just didn't find it important." [Subject #5]

Three: Getting through it

"I would definitely click on 'did I answer correctly', just to know how I did. And I would only click 'what was the correct answer' if the answer was no. 'Book reference'? I don't think I would look at this very often because if it's in a quiz or a test I just want to get it over with. [Subject #9]

Four: Mindful learning

"Yeah, I suppose you wanted to know for your studies. It's useful to know the right answer and you'd use it for, you know, revision and stuff like that." [Subject #1]

"I look for the fact that if I didn't know an answer, why I didn't know an answer. Where was it in the course material that I didn't read about it? The other thing was that, or is a factor, if I happen to answer incorrectly; I want to gain the knowledge. That means I don't know that answer so I want to know... I want to learn." [Subject #6]

“(The lecturer) told us that, don’t worry, it doesn’t count, it’s just to see how much you’ve learned. So when I did that, I was more like learning it. Like I was trying to see oh, did I really fully understand the test itself and so I’m more... I was more motivated to see all those (feedback types).” [Subject #8]

Sense of inquiry

‘Sense of inquiry’ refers to the degree of interest learners hold in an activity. In study two subjects used a variety of terms to describe their sense of inquiry and, again, stated this to be a determining factor when electing to view a feedback type. These terms included ‘like’, ‘curiosity’ and required ‘degree of satisfaction’. This sense of inquiry operated at three different levels. First, the course level,

“(The decision to view a feedback type) depends upon the level of satisfaction that I require out of my course. If I happen to have a desire of learning and getting knowledge, then I would go back to (the elaborative feedback types) ...which will give me confidence to answer that question if it happens again.” [Subject #6]

“I did read it but it was really long and I don’t like ethical stuff.” [Subject #1]

“Some of the questions were linked to my previous educational background and I was really interested, you know, whether I got it right or wrong.” [Subject #7]

Second, at the activity level,

“Just general curiosity. I mean I tend to like (new) things ...I’m always like... if there’s something different like something unique to this programme, I haven’t seen it anywhere else so really driven by that.” [Subject #4]

And finally at the question level,

“As I told you, some questions I found really interesting, some were linked to my previous educational background and some were linked to the interesting aspects of the course.... And I really wanted to know whether I got those right or wrong. Some questions I didn’t really care about but nevertheless I went through all of them just to get the feedback whether I answered them correctly or not.” [Subject #7]

“I think that’s interesting but I only visit that tab (EI2 type feedback) ...I’m really interested in the questions, not every time, probably two/three questions out of ten questions.” [Subject #8]

4.3.4.2 Factor #I2 Response certitude

The ten interviews provided strong support for the hypothesis that learners’ degree of response certitude is a factor affecting whether learners elect to view a feedback type. Seven of the ten subjects stated response certitude to be a significant factor in choosing to move beyond KR and KCR type feedback and view the EX, IB, EI1 and EI2

elaborative feedback types. All seven subjects stated they elected to view the additional feedback when they held a high level of certainty that their answer had been correct. These statements were unprompted, as stated above, explicit reference was not made to any of the potential factors by the interviewer during the interviews.

Three of these subjects referred to a need for proof that the correct answer was indeed correct.

“But if I see the correct answer and I be like no, I’m sure I’m right, I think that’s the correct answer, then I want to prove it so I would see okay book reference where I can find the real correct answer.”
[Subject #8]

“Yeah. Well where did I go wrong? I mean I felt that I was right but where did I go wrong? So if I have an explanation of where did I go wrong and which place did I make my mistake and it could be better.” [Subject #3]

“And if I was... if I disagreed personally or was not sure why this was the right answer, that’s when I went to the third stage which said further explanations actually.” [Subject #7]

To a lesser degree, the subjects stated they would elect to view additional elaborative feedback types when they had a low certainty that their response was correct, when they had guessed.

“Need further explanation, probably for the ones that I just thought... I just didn’t have a clue.”
[Subject #1]

“I think I go to further explanation every time I... when I do wrong or sometimes even I do right, but actually I should’ve guessed it right, like I’m not really sure, 100 per cent sure about it, then I would go to further explanation. So that’s like the quite important for me, information.” [Subject #8]

Study one found that, while, subjects were more likely to view feedback types KCR, IB and EX following an incorrect response, they also elected to view these feedback types following a correct response. The presence of guessing offers one explanation for this.

The model proposed by Kulhavy and Stock (1989) makes two main predictions. First, that response certitude affects the length of time learners spend processing a feedback message. Second, response certitude determines the degree to which feedback affects learners’ performance in subsequent tests.

The model predicts that when subjects hold a high degree of response certitude but respond incorrectly then they will spend a longer time on feedback due to the volume of information processing required,

“In this instance, the successful feedback cycle must eliminate the error response, substitute the correct, and strengthen the new response to the extent necessary for it to appear on the post-test. These multiple activities require considerable effort on the part of the learner. Because a high certitude error is the point of maximum discrepancy, it is also the point at which the model predicts

the largest feedback processing times. The model also predicts that ‘large scale’ elaborative feedback will have the greatest effect when response certitude is low.” (Kulhavy and Stock, 1989: 282)

Study two reveals additional detail relating to why learners spend more time on feedback following an incorrect response. This being that learners may give attention to more information in the feedback message when holding a high degree of response certitude.

Kulhavy and Stock’s (1989) model also predicts that ‘large scale’ elaborative feedback will have the greatest effect when response certitude is low. This prediction was also supported by the study. Subjects stated that having guessed at a question they would be more likely to view the elaborative feedback types beyond KCR.

4.3.4.3 Factor #13 Interpretation of task

The literature review identified learners’ interpretation of the task they face to be a factor potentially affecting whether subjects would elect to view a feedback type. The ten interviews provided support for this hypothesis.

The instructional roles played by multiple choice questions can be organised into two main categories. They serve a testing function by providing a numeric score that allows learners, and their educators, to measure learning. They also facilitate learning by enabling learners to identify gaps in their knowledge and, through the provision of elaborative feedback, assisting learners in closing these gaps.

The learning purpose of the multiple choice questions was emphasised in instructions given to the subjects prior to the start of the activity investigated in studies one and two. Subjects were told that the primary objective of the activity was to revise the course material in order to learn this material more thoroughly. Despite this, while subjects did make reference to both testing and learning, reference to the former was frequent and reference to the latter was infrequent. Six of the ten subjects gave no indication that they recognised the value of the activity in increasing their knowledge of the two topics and did not demonstrate they had actively formed links between the activity and the wider learning objectives of their course.

Subjects generally associated multiple choice questions with testing their knowledge rather than as an opportunity to learn. Four subjects clearly felt that learning the material was a separate activity,

“Just thought oh I know, I’m going to actually read this later, when I revise the exam so I’ll go through this and probably be able to answer it better than I have done.” [Subject #1].

Respondent: “And after that the further explanation. A book reference, yes, to an extent yeah, but I didn’t use this much book reference thing.” [Subject #3]

Interviewer: “Why not?”

Respondent: “Because.... we can have references in the booklet” [Subject #3]

Interviewer: “Yes.”

Respondent: “In the course book... what I feel is I would have to finish the test immediately. And instead of having looking at the booklet things I would look at further explanation.”
[Subject #3]

Interviewer: Okay. So what you’re saying is you wanted to do the test and then do the study afterwards?

Respondent: Yes.

Respondent: Typically it’s just been my habit always that way cos I usually don’t enjoy doing tests unless I either really need to get a sense of much... how comfortable I am with the (material) or I’m (want to) actually test my knowledge. [Subject #4]

Well I’m not an expert on teaching but I would, you know, doubt whether a test is the best way to learn something. [Subject #5]

Subjects interpreted the task as to first complete the activity and second, to do so with as high a score as possible. This orientation can be partly attributed to the design of the activity. In study one, the MCQ questions were not integrated into the online learning activity. Subjects first studied material online in their own time and then answered the multiple choice questions in class. Subjects indicated that their feedback choices may change if the multiple choice questions were more closely integrated into the online activity.

“I think so. I think if I was... if it wasn’t for a test or a quiz and it’s more for like knowledge or training, like a tutorial, then for sure I think after each question” [Subject #4]

“That’s right cos I recall I had to log in from home to do it and even at that point I just did it as a test. I did all the questions and pretty much that’s it. And I think yes, if it’s like a training or... I mean if it’s at any kind of training module then it’s essential I think, it would be great to have something after each question.” [Subject #4]

Learners’ emphasis on the purpose of multiple choice questions as a testing tool was a reason for not electing to view elaborative feedback types beyond KCR. Conversely, those who recognised the role played by multiple choice questions in promoting learning stated this to be a reason to view these feedback types.

“(The lecturer) told us that, don’t worry, it doesn’t count, it’s just to see how much you’ve learned. So when I did that, I was more like learning it. Like I was trying to see oh, did I really fully understand the test itself and so I’m more... I was more motivated to see all those (feedback types). But if, had (the lecturer) had said this is a real exam, it counts, I would have been more like driven by score and aim to get high marks. So it depends.” [Subject #8]

“Respondent: Yeah. I think, exactly, if this were a learning thing I would go through them whereas if it were a testing thing I probably wouldn’t.” [Subject #9]

Subjects’ interpretation of the task provides one explanation why many subjects considered elaborative feedback beyond KCR to be either unnecessary or optional, ‘additional work’.

4.3.4.4 Factor #I4 Level of expertise

The ten interviews provided limited support for the hypothesis that learners’ level of expertise is a factor affecting whether learners elect to view a feedback type.

Butler and Winne (1995) propose two reasons why level of expertise affects learners’ interactions with feedback and potentially determines whether subjects view a feedback type. First, higher level learners have a broader and deeper knowledge of the tactics and strategies required to complete tasks in the subject domain. Second, Butler and Winne state that higher level learners have a broader and deeper knowledge of the tactics and strategies required to complete tasks in the subject domain.

Support for the former was given by one subject who stated that their lack of knowledge relating to a question resulted in a belief that they would not be able to understand the feedback sufficiently.

“For the ones that I just thought I just didn’t have a clue, (I) just didn’t know the theory, you just wouldn’t know so I didn’t know the theory so how am I supposed to know, so I would move on.” [Subject #1]

Related to the issue of learner level is the impact of prior study. Two subjects stated that a lack of preparation would cause them not to view the elaborative feedback types.

“ If I (had) revised I would be more interested in the feedback. Because I didn’t revise that well, I wasn’t really interested in feedback. I just thought oh I know, I’m going to actually read this later, when I revise the exam so I’ll go through this and probably be able to answer it better than I have done.” [Subject #1]

4.3.4.5 *Factor #I5: Performance*

The ten interviews provided support for the hypothesis that learners' performance, that is whether they answer a question correctly, is a factor affecting whether learners elect to view a feedback type.

Study one found subjects were more likely to view feedback types KCR, IB and EX following an incorrect response. As detailed above, under the section 'the existence of a threshold', strong support was given for this finding in the interviews. Nine of the ten subjects explicitly mentioned performance as a factor determining whether they would proceed to view the elaborative feedback types following viewing KR feedback.

"So if I get the right answer then I know what it's talking about and didn't necessarily check everything cos that was happy." [Subject #2]

Unfortunately the format of the interviews did not elicit clarification on whether this behaviour varied across the elaborative feedback types. The focus was on elaborative feedback in general.

4.3.4.6 *Factor #I6 Gender*

No reference was made to gender in any of the interviews. Nor can any discernable pattern emerge from the transcripts that distinguish the male from the female subjects.

4.3.4.7 *Factor #I7 Error type*

This factor was not referred to by the respondents. However, this is explained by the fact that this term describes a concept in the theoretical educational domain and is not likely to be familiar to the subjects. This study does not provide evidence that error type is not a factor potentially affecting students' selection of feedback types but does suggest that students themselves are not consciously aware of this concept.

4.3.4.8 *Factor #E1: Assessment*

Strong support was given for the hypothesis that learners' selection of feedback will change if an activity relates to an assessment. Nine out of the ten respondents declared this to be a factor that would affect whether they chose to view a feedback type. Three distinct scenarios emerged.

First, subjects suggested that their selection of feedback type would change if an activity constituted an assessed piece of coursework. The majority view was that subjects were less likely to view elaborative feedback in this scenario.

"I feel that, like a lot of things where my grade depends on it, I'll do as much as I can to adequately prepare beforehand. But once I actually take it, I'm done. Do you know what I

mean? Like I don't want to see it anymore, like what's done is done, I can't obviously change my answers." [Subject #10]

However, one respondent stated that they were more likely to view elaborative types of feedback if an activity is assessed

"For the final exam I would like to have just the score at the end of the exam. But then say after the exam I go home, I can go back and see what I do right and what I do wrong, like why I did... why I did wrong, that's important." [Subject #8]

The second scenario occurs if an activity constitutes an assessed piece of work and is held in exam type conditions. In this case, there was general agreement among the subjects that receiving feedback as they progressed through an activity would have a negative impact on their performance. Subjects feared that their confidence would be negatively affected by discovering that they had answered a question incorrectly. Subjects stated they would elect to view feedback upon completion of the activity,

"If it is a graded test I think ... getting a response after each question would pose a problem because if you happen... make five incorrect choices in the beginning then... if you know that you're doing something wrong, then it greatly decreases your morale." [Subject #6]

"An exam like this, if I saw halfway through I'm doing bad, you know, getting more and more answers wrong, then for me it would I think only stress me out more and then from there it's just a downward spiral." [Subject #10]

Four subjects mentioned a third scenario in which the activity itself is not assessed but a direct link exists between the activity and a subsequent assessment.

"I think the only time (my choices) would change if this were a practice for a final, people would... I would want to go and look into the further explanation, possibly the book reference." [Subject #9]

4.3.4.9 *Factor #E2 Feedback level*

This factor was not referred to by the respondents. An explanation is that, as described above, all feedback provided in study one is at one level, the task level. In addition, this term describes a concept from the theoretical educational domain and is not likely to be familiar to the subjects.

This concludes the analysis of the nine factors identified during the literature review. As stated above, subjects stated an additional four external factors affected whether they chose to view a feedback type: These factors are introduced and discussed.

4.3.4.10 *Factor #E3: Time limit*

The ten interviews suggest that the presence of a time limit is a factor affecting whether learners elect to view a feedback type.

As described above, measures were taken in the preparation and design of the activity to minimize the impact of a time limit. While the activity generally takes between 25 and 30 minutes to complete, subjects were given 60 minutes. Subjects were told that they had more than sufficient time and also that they were not required to answer all of the questions. Subjects were told that the purpose of the exercise was to learn the material rather than test their knowledge.

Despite these measures, seven of the ten respondents stated that the presence of a time limit was a factor affecting their choice to view the feedback types. Subjects recognized that plenty of time was allowed.

“It was a long time limit but I remember it being like if I read each one I might run out of time. So I was scared about that as far as when there’s a time constraint.” [Subject #10]

“I think we had like almost a whole class time, so we had plenty of time but still I think if I got caught up reading each one I’d probably run out of time, like going through everything.” [Subject #10]

However, the presence of a time constraint still affected subjects’ behaviour. As one subject put it,

“I think it’s like just when you have the time constraint in there, it really shifts your mode because you’re just like okay, I need to get through this now. Because you always get overwhelmed with the content ...” [Subject #4]

Time pressure was also stated as a reason to choose delayed feedback. Three of the ten subjects in study two stated that they had elected to view feedback at the end of the activity and two of these subjects stated a feeling of time pressure to be the reason behind this choice.

“First... I decided which type of feedback I want and I selected the one which gives me the option to have the feedback at the end of the exercise. I did it because I assumed that if I had the feedback given at the end of each question, after answering each question, it would slow me down a little bit because I would start to think about ‘how did I answer?’, ‘why did I answer?’.” [Subject #7]

4.3.4.11 *Factor #E4: Location*

Related to the issue of time is that of location. Seven of the ten respondents stated that their location when taking the activity would affect their selection of feedback. The thoughts of the majority of these subjects are captured in this statement from one interview.

“It’s quite a lot to read... so if you want people to use it as a learning tool people need to be sort of in the learning environment that they like to work in. Some people like to work at midnight, some people like to work at eight in the morning. Some people don’t like to work at the computers here (on campus). Because you might, yeah, this is going to take you two hours, you’re like well I’ll do an hour now and then I’ll watch (television) and I’ll do an hour later after that.” [Subject #1]

Some, but not all, of the respondents indicated that their behaviour would be less task orientated if at home and that they would integrate the study of other course materials into the activity.

Respondent: “If it is a home based test, like one I did yesterday, I open my (course materials) ... I answered the questions through, if I didn’t know about some question then I went back to the course materials and then answered it....

.....So in the process of finding it out, I had to read at least one page (from the online materials) to find (the answer to) one question. In the process, I had read everything. “ [Subject #6]

Interviewer: “So you’re reading the original materials before answering (a question)?”

Respondent: “Yes. If it is home based test.” [Subject #6]

4.3.4.12 *Factor #E5: Question focus*

The ten interviews suggest that the focus of a question, whether a question tests recall or application of the material, affected subjects’ choice to view a feedback type. This is termed ‘question focus’ in this research.

Three of the ten subjects drew a distinction between questions that test the recall of course concepts and questions that require the application of these concepts in novel scenarios. These subjects stated that their choice of feedback would change between the two questions types.

“Every time... I want(ed) to know if I did it right or wrong. And if (right) ... (and) it’s just memory question, then I will just go to next question instead of trying to see oh why I did right. Because it’s just memory questions. But if it’s no, then I will definitely go to ‘what is the correct answer’.” [Subject #8]

“But if it was something (like) accounting which is not exactly... then I believe further explanation would be necessary. Because I would know which part of the book it came from. But how to apply that knowledge into the question, I need to know that. Like if I have further explanation that would be explaining ... the question.” [Subject #6]

The issue of question focus does appear in prior research. Two studies identified in the literature review investigated the relative effects of feedback on questions that test the recall of facts and those that test inference. Peeck (1979) posed a series of fact and inference questions to fifth and sixth grade elementary schools children following their study of a 900 word text. An analysis of variance found that the provision of feedback had a stronger positive effect on subsequent performance for inference questions than for fact questions. However this only held for elaborative types of feedback. On the other hand Andre and Thieman (1988) found that the provision of feedback to factual questions had a greater effect than for inference questions.

4.3.4.13 Factor #E6: Subject matter

The ten interviews also suggest that subjects’ choice to view a feedback type would be affected by whether a question covers quantitative or qualitative material. This is termed ‘subject matter’ in this research.

Subjects had studied multiple choice questions courses in accounting, mathematics and economics during the term that preceded the interviews. Seven of the ten subjects stated that their selection of feedback type would change according the subject being studied and made explicit reference to their other courses.

“I think if it were accounting or math the book reference wouldn’t be so important. It wouldn’t be very applicable. The further explanation would be quite important and the examples would be important... I think because it’s a bit more mathematical and it’s a bit more difficult to grasp whereas with informatics if you see the correct answer ... you might say oh I don’t know why I didn’t think of that.” [Subject #9]

“For the pre-course ones we had, the accounting and the maths, before we started that I think (the elaborative feedback types) might’ve been really helpful. I know it’s a bit more difficult because that’s more quantitative stuff but a lot more description can enhance the learning. In those courses specifically I think would be really good.” [Subject #10]

“If it was a finance thing and I don’t know how to solve the problem. Then I think further explanation... further explanation and examples would do the trick.” [Subject #6]

One subject stated the reason to be that quantitative questions required a sequence of steps, a process, to answer.

“If it’s accounting, I don’t need to see what’s the correct answer because I know that’s the correct answer, but I want to see the process. I want to know, even if I do it right, I want to still know the process because sometimes it’s just because I’m lucky, because the process probably is wrong but the answer is right, then so, then I wouldn’t learn anything from doing that question.” [Subject #8]

4.3.4.14 Summary of question two

Support for the nine potential factors identified in the literature review varied. The interviews provided no evidence that the following three factors had affected whether subjects chose to view the feedback types.

#I6 Gender

#I7 Error type

#E2 Feedback level.

Some support was given for factor ‘#I4 level of expertise’ however this was not sufficient for clear conclusions to be reached. However, strong support was provided for the following five factors:

Factor #I1 Learner expectations

Factor #I2 Response certitude

Factor #I3 Interpretation of task

Factor #I5 Performance

Factor #E1 Assessment

Four additional potential external factors emerged during the analysis. These being the presence of a time limit, the location at which the activity is conducted, the subject matter being studied and whether a question tests recall or application of the original material. These were termed factor ‘#E3 time limit’, factor ‘#E4 location’, factor ‘#E5 question focus’ and factor ‘#E6 subject matter’. That all four factors affected whether subjects chose to view a feedback type does appeal to common sense. It can be assumed that subjects’ behaviour will change if a time limit is removed and also if they are located in a less formal situation than the classroom. It can also be assumed that the nature of a question, its focus and the subject matter it refers to, also affects the type of feedback learners require to achieve closure.

4.3.5 Further findings

Seven further findings emerged during the analysis relating to questions one and two. These are now introduced and discussed.

Further findings 1 of 7: Learners elect to view the minimum volume of TBI feedback required to achieve closure

Feedback load, that is the volume of feedback presented was a concern for the subjects. Seven of the ten subjects made explicit reference to load during their interviews. Subjects in this study indicated they pursued routes to achieving closure that involved viewing the smallest volume of feedback material. Subjects generally held negative views towards high feedback loads and gave the following three reasons for this.

Reason one: Requirement for time efficiency:

Seven of the ten subjects made reference to time efficiency during their interviews. Subjects considered time to be a precious resource and were keen to minimize the time spent on feedback. Subjects wanted to access the information required from feedback as fast as possible.

“Because ... to be honest students, they want quick access to information, immediate access to information.” [Subject #3]

One example of this strategy is subjects’ comparison of the IB and EX feedback types. Both these types provide information that contains the knowledge required to answer the question correctly. IB type feedback presents the relevant section of the instructional material and EX provides explanations as to why each answer is correct or incorrect. Hence the two feedback types can be considered interchangeable. However, in study one subjects exhibited a slight preference for EX type feedback. Subjects viewed EX for 8.89 per cent and IB for 6.55 per cent of the questions they answered. In this series of interviews, while subjects made positive statements in relation to the usefulness of IB type feedback, they reiterated a preference for EX type feedback and also stated a consistent rationale. Five of the ten subjects stated that EX type feedback was the more time efficient choice.

“Because I felt I won’t be going back and looking at... the book. I want to know it immediately, like where did I go wrong immediately, once the question is answered. So that’s why I vote for further explanation which is a short gist of the information. So I immediately know that where did I go wrong.” [Subject #3]

“(Further explanation) This might be more useful than the book reference because people would like to go straight into the explanation rather than reading a whole section of the book.” [Subject #9]

“Instead of having looking at the book reference I would look at further explanation. ... I want to know it immediately, like where did I go wrong immediately, once the question is answered.” [Subject #3]

The format of the activity could explain the importance subjects ascribed to being efficient with their time. Although subjects were given 60 minutes to complete an activity designed to take between 25 and 30 minutes to complete, were told explicitly that they had been given a surplus of time to complete the activity

and that they did not need to complete all of the questions, the mere presence of a limit may have made subjects more conscious of time than they would have been otherwise.

“Because that was a test as I said. It was... I perceived it as a test and normally tests are time limited, you kind of have that pressure that you have to finish it in time, go through the... whole lot as quick as possible but obviously think about the answers and the questions. So it’s probably the time pressure.” [Subject #5]

Despite this, the impression given by the interviews is that subjects generally value time efficiency and that this is not dependent on the presence of a time limit to an activity.

Reason two: A high feedback load dampens a sense of progress

Two subjects stated that a high feedback load dampened a sense of progress through the activity.

“I mean my own say (sic) would be.... I wouldn’t want to get bogged down into each question and like looking, okay, well okay good I got this wrong, let’s figure out exactly what happened before moving to the next.” [Subject #4]

“You just lose your flow of doing a test. “ [Subject #1]

To these subjects, successfully reaching the end the activity seemed to be of greater importance than closing knowledge gaps relating to the questions.

Both this issue, and subjects’ requirement for time efficiency, can be explained by subjects’ interpretation of the task. The subjects in study two generally associated multiple choice questions with testing their knowledge rather than as an opportunity to learn. Subjects interpreted the task as to first complete the activity and second, to do so with as high a score as possible.

Reason three: Subjects found higher volumes of feedback difficult to absorb

Seven out of the ten subjects stated that higher volumes of feedback, that is higher volumes of feedback than they felt to be required, are difficult to absorb.

“Just too much at once because it’s kind of like overwhelming. You probably want to look (at book reference or things like further explanation) ... if you were like studying... But sometimes there’s like a lot of questions where it’s just right or wrong and you want like a brief reason as to why.” [Subject #4]

“They are, you know, if it’s too much written, it just looks busy... But if it’s in one line for instance, a highlight sort of bold or anything, it gets your eyes, catch the phrases, what they want to bring out of it.” [Subject #5]

Cognitive load theory provides a theoretical explanation for these statements on feedback load. Cognitive load theory, discussed in the section ‘Theoretical framework’ (See page 65), emerged from cognitive psychology and focuses on the role played by a limited working memory in processing information. It is not possible to look at prior research to investigate this further. The literature review did not identify prior research into TBI feedback that considered the relationship between of the load of feedback and cognitive load.

That learners find higher volumes of feedback problematic poses a problem to designers of instructional materials. In studies one and two, the occasions on which subjects elected to view most feedback appeared to be those that most resulted in facilitating the deep learning orientated behaviour that is the aim of most educators. However, there seems to be a tension between providing sufficient feedback to promote such experiences and satisfying learners’ general preference for low feedback load.

Further findings 2 of 7: Why learners prefer immediate feedback

Study one found that 137 of the 146 subjects preferred to receive feedback immediately after responding to a question as opposed to after answering all the questions. In study two, three students stated that they had elected to view delayed feedback and all stated that their reason for doing so was a concern regarding the time required to read the feedback. This is discussed further in the analysis of question two above under ‘Factor #E3: Time limit’ (See page 164).

Six subjects stated they chose to view feedback after each question. Two of these subjects described their rationale using general terms such as ‘prefer’, ‘want’ and ‘better’. However, three subjects were able to elaborate further stating that viewing feedback immediately was the more efficient approach. These subjects felt that receiving feedback immediately after responding to a question was the efficient choice as they would already be holding the detail of the question in mind. Receiving feedback at the end of the exercise would require them to restudy a question in order to recall the thought behind their response and therefore required more effort.

“I would choose after each question... because at the end of the whole exercise you might forget what the question is and why you chose a specific answer.” [Subject #9]

“I just went to click this one, 'after each question' because if I get a feedback after a question itself then it helps me remember the ... correct response for my incorrect answer. That helps me relate to the question better because I have the memory. But if I have a look at the end of the exercise, then I have to go back to the question again, then ... understand that question, what it is about, look at what I incorrectly answered, then go back to ... what was the correct answer and makes the process long.” [Subject #6]

“I always choose after each question because for me it’s very important to get instant feedback from each question so that I don’t forget. ...If I get all the answers or feedback at the end of the exercise I

have to go back and I have to rethink why I chose this instead of the right answer. It's not as efficient and ... it takes double time." [Subject #8]

Further findings 3 of 7: Support for the existence of a threshold

The literature review identified a body of research proposing the existence of a threshold, at the level of KR plus KCR type feedback, beyond which the study of feedback becomes either ineffective or inefficient (Phye 1979; Kulhavy, White et al. 1985; Dempsey and Litchfield 1993). These studies propose that feedback provided beyond this threshold is ineffective or inefficient.

Study one supported the existence of a threshold, at the level of KR plus KCR type feedback. Feedback beyond these types was viewed significantly less frequently. The subsequent analysis questioned whether the benefits of elaborative feedback types beyond KCR justified the development cost.

These interviews were able to illuminate aspects of subjects' own perceptions of this threshold. Subjects revealed a process in which they moved from KR type feedback to KCR and then, on specific occasions, to the remaining types of elaborative feedback. This process is now described.

Learners first action is to view KR type feedback

All subjects stated that their first action was to view KR type feedback. The purpose being to seek the affirmation described above. The subjects wanted to know whether they had responded correctly.

"Every time I would click... I clicked on the first option (KR) because it's... I want to know if I did it right or wrong." [Subject #8]

"We wanted to know... which one was right because...you wanted to know, you know." [Subject #1].

Subjects considered the wish to know whether they had answered a question correctly or incorrectly a matter of course,

"Of course you want to know the answers." [Subject #2]

When learners respond correctly, KR is generally sufficient to provide a sense of closure

Study one found that subjects were less likely to view elaborative feedback types KCR, IB and EX after responding to a question correctly. This finding was supported by the interviews. Nine of the ten subjects stated KR type to be sufficient when they had responded correctly.

"So if I get the right answer then I know what it's talking about and didn't necessarily check everything cos that was happy." [Subject #2]

When learners respond incorrectly, KCR type feedback is generally sufficient to achieve closure

After responding incorrectly, all subjects stated that they would elect to view KCR type feedback to discover the correct answer after viewing KR type feedback. This accurately reflects the quantitative data collected in study one.

Study one found a noticeable difference between the frequency at which subjects elected to view KCR type feedback and the frequency at which subjects elected to view the other forms of elaborate feedback thus providing support for the existence of a feedback threshold. This indicates that viewing KCR type feedback is generally sufficient to achieve closure when subjects responded incorrectly to a question. During this series of interviews, subjects gave three reasons for this.

First, discovering the correct answer was sufficient to close the gap between what the subject understood and what needed to be understood to answer the question correctly.

“So when I do it wrong, then I go to... I want to know the right answer and if I see the right answer and I’d be like, okay I understand.” [Subject #8]

“You know, upon seeing the right answer then you remember why that was the right answer I just instantly knew that, you know, okay I made a mistake because this makes sense.” [Subject #7]

Three subjects referred to instances in which they had narrowed the choice of options to two. If, after responding to a question incorrectly, they discovered that the alternative response was correct then they did not require further information to achieve closure on the question.

“I had the two options and if I was wrong, it was the other option always. The correct answer was there, so (it was a bit annoying but) I figured it out eventually.” [Subject #5]

Second, while subjects were highly motivated to discover whether they had answered correctly and also to view the correct answer if they did not, the need for additional explanation was seen as less of a requirement.

“Personally, (I) don’t think I would read further explanation but then that’s just... I think it’s like added interest.” [Subject #4]

Respondent: “Why did I move on? I just didn’t find it important.” [Subject #5]

Interviewer: “Knowing the correct answer, that’s important?”

Respondent: “That’s right.” [Subject #5]

Interviewer: “But knowing an explanation of that is not important?”

Respondent: “No.” [Subject #5]

Respondent: “Because I want to know whether the answer’s correct or not. And if it wasn’t, if it says no, I want to know which one and that was it.” [Subject #5]

Thus the achievement of closure did not always require a comprehensive resolution of the gap between what was understood by a subject and what had been required to be understood in order to have responded correctly. The degree to which this gap must be resolved in order to achieve closure varied.

The third reason stated for KCR type feedback being sufficient to achieve closure was a negative perception of high volumes of feedback. This issue of feedback load is discussed below under ‘feedback load’.

The concept of a threshold, at KCR level, beyond which feedback becomes either ineffective or inefficient was supported by this series of interviews.

Further findings 4 of 7: Why learners move beyond KCR threshold

However, the data gathered in study one showed that any threshold was not absolute in that subjects did view feedback types IB, EX, EI1 and EI2.. In this series of interviews subjects gave three specific reasons for the instances in which they did elect to view the additional elaborative feedback types. All three were factors identified in the review of prior research;

Factor #I1 learner expectations,

Factor #I2 response certitude

Factor #I3 interpretation of task

The role played by these factors has been discussed in more detail in the analysis of question two above. A minority of subjects stated factors ‘#I1 learner expectation’ and ‘#I3 interpretation of task’ to affect whether they elected to view elaborative feedback beyond KCR type feedback following an incorrect response. However, as discussed above, seven of the ten subjects stated response certitude to be a significant factor in choosing to move beyond KCR type feedback and view the EX, IB, EI1 and EI2 elaborative feedback types.

As stated in the analysis of question one above, subjects' purpose in electing to view the elaborative feedback types was to resolve gaps in their knowledge. Specifically, to understand or to learn.

“I didn’t always click on it (IB Type feedback). but ... when a correct answer didn’t fully explain it to me. Like, so I got a question wrong, now this is the right answer but why is this is the right answer and that’s when I go to the book.” [Subject #10]

“(If) I’m not so sure... between two... answers then I need the full explanation and ... so ... I normally go here, (Elaborative feedback types beyond KCR).” [Subject #2]

In this series of interviews subjects explicitly stated that these three factors affected their choice to view the elaborative feedback types beyond KCR.

In addition, subjects stated a further seven factors had affected or may affect their decision to view feedback in general. These being factors ‘#I4 level of expertise’, ‘#I5 Performance’, ‘#E1 Assessment’, ‘#E3 time limit’, ‘#E4 location’, ‘#E5 question focus’ and ‘#E6 subject matter’. This suggests that these seven factors also affect subjects’ choice to view the elaborative feedback types beyond KCR.

Further findings 5 of 7: Learners’ preference for elaborative feedback beyond KCR that focuses specifically on the knowledge gap

A hypothesis emerged in study one that learners hold a preference for elaborative feedback beyond KCR that focuses specifically on the knowledge gap, that is the gap between what was understood and what had been required to be understood in order to have answered a question correctly.

Some support for this hypothesis emerged during this series of interviews. EX type feedback was seen as being effective in closing any knowledge gap that remained after subjects had viewed KR and KCR type feedback.

“Instead of having looking at the book reference I would look at further explanation. ... I want to know it immediately, like where did I go wrong immediately, once the question is answered.” [Subject #3]

IB was also seen to being able to close this gap but was considered less efficient.

“But if I see the correct answer and I be like no, I’m sure I’m right, I think that’s the correct answer, then I want to prove it so I would see okay book reference where I can find the real correct answer.” [Subject #8]

“(Further explanation) This might be more useful than the book reference cos people would like to go straight into the explanation rather than reading a whole section of the book.” [Subject #9]

“Instead of having looking at the book reference I would look at further explanation. ... I want to know it immediately, like where did I go wrong immediately, once the question is answered.” [Subject #3]

However, subjects were not vocal regarding the role played by the extra-instructional feedback types EI1 and E12. These feedback types were not generally seen as being core to the task. Some students considered these feedback types to be unnecessary or ‘nice to have’.

“And ‘further examples’ is always nice. But once again I think it would depend on the level of depth required for the topic.” [Subject #4]

“From my point of view ‘further examples’ is more like, it’s like further information isn’t it?” [Subject #8]

“‘Explanation of terms’. That’s helpful so that I don’t have to Google it. But even if it’s not there, it’s quite easy to just Google it. So it’s, for me if it’s there, great, it saves time. But if it doesn’t, like I wouldn’t complain.” [Subject #8]

Subjects placed higher value on feedback focusing specifically on the knowledge gap.

Further findings 6 of 7: Explanations for the variety in student behaviour

Study one plotted subjects’ α values for feedback types KCR, IB, EX, EI1 and EI2 and found that, while the distributions contained a marked positive skew, subjects’ α values spread over a wide range. In other words, while the majority of students elected to view the elaborative feedback types at a similar frequency, a minority of subjects elected to view these feedback types at noticeably different frequencies.

Differences in performance largely explained this variation in behaviour, particularly for feedback types KCR, IB and EX. However, difference in performance did not explain all the variation and other factors must be present. This series of interviews provides support for the fact that a further nine of the factors identified affected subjects’ choice to view the feedback types. Four of these are classified as external factors and are not specific to a particular learner (‘#E1 Assessment’, ‘#E3 time limit’, ‘#E4 location’, ‘#E5 question focus’ and ‘#E6 subject matter’). This leaves four factors that are able to explain the remaining variety in student behaviour; factors ‘#I1 learner expectations’, ‘#I2 response certitude’, ‘#I3 interpretation of task’ and ‘#I4 level of expertise’.

It can be assumed that the degree of certitude a learner holds in their responses will vary across the questions in the exercise and that also that differences in response certitude across exercise as a whole will largely derive from the other three factors. Thus we can conclude that it is the three remaining factors that account for the variation in the distribution of subjects’ α values not explained by differences in performance:

Factor #I1 learner expectations

Factor #I3 interpretation of task

Factor #I4 level of expertise

Further findings 7 of 7: Learners require feedback at the question level to achieve closure

One further and, perhaps, less notable finding was a stated preference for feedback at the question level. Eight of the ten subjects stated that knowing their overall score alone was insufficient. To articulate this point, one respondent [Subject #7] referred to dissatisfaction with a previous set of tests, from a different course, in which feedback at the question level was not provided.

“I found (the test) ... good ... it really forced us to read the case study, but there (was) no feedback at all, and we ... didn’t know why we did wrong or right, that was... really annoying, without feedback.” [Subject #7]

4.3.6 Summary and conclusions

The aim of study two was to investigate subjects’ reasoning behind their choices made in study one. Specifically, the series of interviews aimed to address the following two questions,

1. Why do learners elect to view TBI feedback?
2. What factors affect whether learners elect to view the TBI feedback types?

4.3.6.1 Why do learners elect to view TBI feedback?

This study found that subjects’ overarching reason for viewing feedback was to achieve closure on the task. When viewing the different feedback types, subjects pursued closure first by seeking affirmation and second by resolving ‘knowledge gaps’, that is any gap between what the subject understood and what had needed to be understood in order to have responded to a question correctly. When resolving knowledge gaps, most subjects stated that their primary purpose was a need to understand. To a lesser extent, subjects stated a requirement to learn.

As discussed in the section ‘Constructivism’ under ‘Theoretical framework’ (See page 65) , the role of feedback in addressing knowledge gaps has appeared in prior research (also 1989: 121, Hattie and Timperley’s 2007: 82). In much constructivist thought it is the identification and resolution of knowledge gaps that leads to the occurrence of learning. Ernst Von Glasersfeld (1989: 136) uses the term ‘perturbation’ to describe the state when a knowledge gap is identified and ‘equilibrium’ to describe the state at which closure is achieved regarding this perturbation. The role of feedback in promoting learning can therefore be considered in terms of stimulating a perturbation and assisting in a return to equilibrium.

4.3.6.2 What factors affect whether learners elect to view the TBI feedback types?

This study then investigated the nine factors identified in the literature review that potentially affect whether learners choose to view a feedback type. Support was provided for factors ‘#I1 learner expectations’, ‘#I2 response certitude’, ‘#I3 interpretation of task’, ‘#I5 performance’ and ‘#E1 assessment’.

Factor ‘#I5 performance’ was found to be the most significant. However findings relating to three further factors are of note. First, the ten interviews provided support for the hypothesis that factor ‘#I2 response certitude’, the degree to which learners believe their response to be correct, is a factor affecting whether learners elect to view a feedback type. Seven of the ten subjects stated response certitude to be a significant factor in choosing to move beyond KR and KCR type feedback and view the EX, IB, EI1 and EI2 elaborative feedback types. A key prediction of Kulhavy and Stock’s (1989) model of response certitude is that subjects spend a longer time on feedback when they hold a high degree of response certitude that their response had been correct but in fact

responded incorrectly. Study two suggests a previously unknown detail relating to this phenomenon. Learners give attention to more information in the feedback message when holding a high degree of response certitude.

Also of note were subjects' statements relating to factor '#I3, interpretation of task'. Subjects indicated strong performance related behaviour, and stated this to be a reason not to view the elaborative feedback types beyond KCR. As discussed above, this questions the case for providing elaborative feedback.

Support was given for the hypothesis that factor '#E1 assessment', whether an activity is assessed, affects learners' selection of feedback. Nine out of the ten subjects declared this to be a factor that would affect whether they chose to view a feedback type. Three distinct scenarios emerged. First, if the activity constituted an assessed piece of coursework. Second if the activity constituted an assessed piece of work and was held in exam type conditions and third, if the activity itself is not assessed but a direct link exists between the activity and a subsequent assessment.

Four additional potential factors emerged during the analysis. These being the presence of a time limit, the location at which the activity is conducted, the subject matter being studied and whether a question tests recall or application of the original material. That all four of these factors affected whether subjects chose to view a feedback type does appeal to common sense. It can be imagined that learners' behaviour will change if a time limit is removed and also if they are located in a less formal situation than the classroom. It can also be imagined that the nature of a question, its focus and the subject matter it refers to, also affects the format of feedback learners require to achieve closure.

These four factors were added to the list of factors considered for investigation in studies three and four.

4.3.6.3 Summary of further findings

Further findings were organised under seven headings. First, Feedback load, that is the volume of feedback presented was a concern for the subjects. Seven of the ten subjects made explicit reference to load during their interviews. Subjects sought to view the minimum load of TBI feedback required to achieve closure and gave three reasons for this. First, a need to be efficient with their time. Second, that a high feedback load dampens a sense of progress through the activity. Third, subjects found higher volumes 'overwhelming'.

Efficiency, albeit of a different nature, was also stated as the reason behind subjects' preference for immediate rather than delayed feedback. Study one found that 137 of the 146 subjects preferred to receive feedback immediately after responding to a question as opposed to after answering all the questions. In this study the majority of subjects stating they had selected immediate feedback did so because they considered immediate feedback to be more efficient. This being because they would still be holding the detail of a question in mind when viewing the feedback message.

The quantitative data collected in study one provided support for the existence of a feedback threshold at the level of KR plus KCR type feedback beyond which the study of feedback becomes either ineffective or inefficient. Study two provided some support for the existence of such a threshold and also detail regarding subjects' perception of

this threshold. However, this study also suggests that the instances in which subjects moved beyond KCR may be of particular importance to promoting learning of the material.

A hypothesis emerged in the analysis of study one that learners hold a preference for elaborative TBI feedback types that focus on any 'knowledge gap'. In other words, they will elect to view these feedback types more frequently. Some support for this hypothesis emerged during this series of interviews. Subjects did view the elaborative feedback types in order to resolve knowledge gaps. EX type feedback was seen as being effective in closing any knowledge gap that remained after subjects had viewed KR and KCR type feedback. IB was also seen to being able to close this gap but was considered less efficient. The extra-instructional feedback types EI1 and EI2 were not generally considered to be core to the task. Some students considered these feedback types to be unnecessary or 'nice to have'.

This study also provided explanations for unexplained variance in the frequency at which subjects elected to view the TBI feedback types in study one. When subjects' α values for feedback types KCR, IB, EX, EI1 and EI2 were plotted on a graph, it was found that subjects elected to view these elaborative feedback types at varying frequencies. Differences in performance largely explained this variation but not completely. This series of interviews suggests that factors '#I1 learner expectations', '#I3 interpretation of task' and '#I4 level of expertise' account for the additional variation.

4.3.6.4 Conclusions

This series of interviews provided explanations for many of the conclusions reached in study one. In particular for the existence of a feedback threshold, for the instances in which subjects cross this threshold and for subjects' preference for immediate feedback. Some conclusions, such as the role played by response certitude, are supported by prior research. Others, such as the impact of question focus, offer avenues of inquiry in studies three and four.

Perhaps the findings of this study of most note are those relating to why learners elect to view feedback and factors that affect whether they choose to view the different TBI feedback types. Learners view feedback in order to achieve closure on the task. When presented with different types of TBI feedback, learners pursue closure first by seeking affirmation and second by resolving knowledge gaps. Learners resolve these gaps primarily to achieve understanding and secondly through a requirement to learn.

This conclusion can be combined with the additional findings discussed above to produce the basis for a model of the process through which learners engage with the TBI feedback types. This study suggests that, when presented with a choice of TBI feedback types, learners first elect to view KR type feedback in order to receive affirmation that their response was correct. Following this learners proceed to view the elaborative feedback types if a knowledge gap has been identified. Learners seek the most efficient means of resolving this gap, that is they seek to view the minimum load of TBI feedback required. KCR type feedback is often sufficient to close this gap and is therefore the most efficient first choice. Following this, subjects view the remaining feedback types in the order of their specificity on the knowledge gap. First EX type feedback, then IB and then the extra-instructional feedback

types EI1 and EI2. This model of interaction with the different feedback types is developed and discussed in more detail in the 'Discussion' section of this thesis (See page 260).

To educators, the most pertinent aspect of this model is likely that relating to knowledge gaps. In constructivist thought it is the identification and resolution of knowledge gaps that leads to learning. Ernst Von Glasersfeld (1989: 136) uses the term 'perturbation' to describe the state when a knowledge gap is identified and 'equilibrium' to describe the state at which closure is achieved regarding this perturbation.

To educators it would be heartening to discover that learners seek to comprehensively resolve such gaps when they are identified. However, this study suggests that this is not always the case. The achievement of closure did not always require a comprehensive resolution of the gap between what was understood by a learner and what had been required to be understood in order to have responded correctly.

The extent to which a perturbation needs to be resolved in order for learners to achieve closure depends on the factors discussed above. Therefore the level of learning that occurs during such activities also depends, to a degree, on these factors.

4.4 STUDY THREE: THE IMPACT OF SUBSEQUENT ASSESSMENT AND QUESTION FOCUS (2009)

4.4.1 Introduction

This study comprised a technology based observation of 11 cohorts of students studying one or more of three different items of TBI. This study re-examined the hypotheses explored in study one and investigated additional hypotheses that arose during the analyses of studies one and two. This study took place between July and December 2009.

In structure, this study followed the same design as study one. The study included the data from study one plus additional data from 10 cohorts of students studying one or both of two further items of TBI. These being the “Accounting Primer” and “Pre-study Mathematics” courses described in the section “The three items of Technology Based Instruction (TBI)” (See page 102). These two courses were made available to subjects for the two months immediately prior to their arrival at the host institution to begin their studies. Subjects studied both courses at a time and place of their choosing.

As in study one, subjects were presented with six feedback options after responding to a multiple choice question. Each feedback item corresponding to one of the six feedback types in the taxonomy of TBI feedback. Subjects were also given the choice either to view feedback immediately following their response to a question or after having responded to all of the questions. As before, the software tool recorded all ‘clicks’ made by subjects in the software user interface.

The hypotheses for this study were,

1. Learners will elect to view feedback more frequently than they do not
2. Learners will elect to view KCR and KR type feedback more frequently than they elect to view either EX, IB or EI type feedback
 - 2.1. For elaborative feedback type beyond KCR: Learners will elect to view EX most frequently followed by IB and then EI type feedback.
 - 2.2. The distribution of α values for elaborative feedback types will contain a marked positive skew
3. Learners will elect to view immediate feedback more frequently than they elect to view delayed feedback
4. Learners’ choice to view a feedback type is dependent on:
 - 4.1. Whether a question is answered correctly or incorrectly (Factor #I5 Performance)
 - 4.2. Whether a question tests recall or application of the study material (Factor #E5 Question focus)

- 4.3. Whether the material being studied will be formally assessed (Factor #E1 Assessment)
5. The majority of learners will revisit a previously viewed question
6. Learners' use of feedback will not change significantly over the course of the activity

4.4.2 Formation of hypotheses

The approach to forming hypotheses for study three began by adopting the hypotheses adopted in study one. These were,

1. Learners will elect to view feedback more frequently than they do not
2. Learners will elect to view KCR and KR type feedback more frequently than they elect to view either EX, IB or EI type feedback
3. Learners will elect to view immediate feedback more frequently than they elect to view delayed feedback
4. Learners' choice to view a feedback type is dependent on:
 - a. Whether a question is answered correctly or incorrectly (Factor #I51 Performance)
 - b. Gender (Factor #I6 Gender)

Next additional hypotheses were incorporated based on the results of study one which found that subjects viewed EX most frequently followed by IB, EI1 and then EI2 type feedback. Also, that the distribution of α values for elaborative feedback types contained a marked positive skew. These two findings were included in study three as the following hypotheses.

- 2.1. For elaborative feedback types beyond KCR: Learners will elect to view EX most frequently followed by IB and then EI type feedback.
- 2.2. The distribution of α values for elaborative feedback types will contain a marked positive skew

Following this, hypotheses were added based on the results of study two. Study two identified an additional four factors that potentially affect whether a learner elects to view a feedback type resulting in a total of 13 possible factors, shown in table 23, for consideration for investigation in study three.

Table 23: The 13 possible factors for consideration for investigation in study three

Internal factors	External factors
Factor #I1 Learner expectations	Factor #E1 Assessment
Factor #I2 Response certitude	Factor #E2 Feedback level
Factors #I3 Interpretation of task	Factor #E3 Time limit
Factor #I4 Level of expertise	Factor #E4 Location
Factor #I5 Performance	Factor #E5 Question focus
Factor #I6 Gender	Factor #E6 Subject matter
Factor #I7 Error type	

Three of these factors were chosen for investigation in study three through the following process of elimination.

Factors #I1 Learners' expectations, #I2 response certitude, #I3 interpretation of task, factor #I8 error type and #E2 feedback level

These factors were eliminated from study three for the same reasons detailed in the report of study one (See page 126).

Factors #E3 time limit and #E4 location

Factors '#E3 time limit' and '#E4 location' were eliminated on practical grounds. The materials used in study one took place in a computer laboratory and students were set a time limit within which to complete the activity. However, the host institution cancelled the use of these materials in 2009. All suitable TBI materials in 2009 were studied off campus and subjects were not set a time limit. Thus it was not possible to obtain data in which a single item of TBI was used by equivalent groups with different time limits or at different locations.

Factor #I4 Level of expertise

Level of expertise was eliminated from study three as data that could be used to measure this factor, such as course scores and overall degree scores, were not available for up to 18 months following subjects' use of the TBI materials. For this reason, investigation of factor '#I4 Level of expertise' was postponed until study four.

Factor #E6 Subject matter

An analysis of the questions across the three items of TBI revealed that all of the questions in the Ethics and Strategy Exercise were qualitative while all of the questions in the Pre- Study Mathematics course were quantitative. Further analysis revealed that all of the quantitative questions across the tests tested application of the material. These results are shown in table 24 below,

Table 24: Question focus and subject matter of question in the three items of TBI

	Question focus		Subject Matter	
	Recall	App	Quantitative	Qualitative
* Accounting Primer Course	13	5	4	14
Pre-Study Mathematics		25	25	
Ethics and Strategy exercise	24	25		50

* - All quantitative questions test application, only one qualitative question tests application.

Therefore it would not be possible to separate the effects of the subject matter of a question from that of whether the question tested recall or application of the material. It was necessary to choose either #E5 question focus or #E6 subject matter and the latter was eliminated as a factor for investigation.

Factor #E17 Gender

It was not possible to include gender in study three as this was not known for the majority of subjects.

This resulted in factors '#I5 performance', '#E1assessment' and '#E5 question focus' forming the factors for investigation under hypothesis four of study three.

4.4.3 Design

This study adopts the technology based observation approach described in the section 'Research methods' (See page 85) and described further in the section 'Design' in the report of study one (See page 128). However the selection of TBI, the factor constructs adopted in hypothesis four, the selection of subjects and issues relating to reliability and validity do differ. These topics are now introduced and described.

4.4.3.1 Factor constructs

Consideration of factor '#E1 Assessment'

Study two revealed three different scenarios in which assessment potentially affects whether subjects choose to view different feedback types within a TBI activity. First, if the TBI activity constitutes an assessed piece of coursework. Second, if the TBI activity constitutes an assessed piece of work and is held in exam type conditions. Third, if the activity itself is not assessed but a direct link exists between the activity and a subsequent assessment.

None of the multiple choice questions within the three items of TBI for which data were collected were assessed either as coursework or as part of an assessed exam. However, some student cohorts were assessed on the "Accounting Primer" and the "Pre-study Mathematics" courses via a formal exam upon their arrival at the host

institution. In study three, factor ‘#E1 Assessment’ refers to this subsequent assessment. In other words, to whether students were subsequently assessed on the TBI material.

Consideration of factor ‘#E5 Question focus’

From study two, a hypothesis emerged that subjects’ choice to view a feedback type is determined by whether a question tests recall or application of the material. This is termed factor ‘#E5 Question focus’ in this research.

In order to investigate this factor, it was necessary to flag each question in the TBI activity. Academic staff responsible for the Accounting Primer course and the Pre-study Mathematics course classified each question as testing either recall or application of the course material.

Selection of the TBI material

Rationales for the selection of the three items of TBI included in this series of studies are presented in the section ‘The three items of Technology Based Instruction (TBI)’ (See page 102). All three items of TBI were included in study three for two reasons. First, to secure as much relevant data as was possible. Second, to ensure that a sufficient volume of data relating to both recall and application questions was included.

4.4.3.2 Selection of subjects

The strategy for the selection of subjects was to include data from as many subjects as possible in the initial analysis and subsequently investigate sub groups of these data as required. The starting point was to gather data for all subjects for whom the online system had collected data as of December 2009. These data were then subject to the processing detailed in the earlier section ‘Analysis of quantitative data: Data processing’ (See page 93).

Summary data relating to the resulting cohorts are presented in table 25 below. Data relating to individual students were not captured for the ‘Pre-Study Mathematics’ tests due to an error in the reporting and these data were included in the study as anonymous data. These anonymous submissions are shown in brackets.

Table 25: Summary data relating to the cohorts in study three

Test	Year	Cohort	Number of submissions	No. of questions answered	Subsequently assessed?
Ethics and Strategy exercise	2008	MSc Management	146	8,103	
Accounting Primer Course	2009	BEST Programme	19	281	
		Full time MBA	58	1,262	
		IBSc Programme	51	981	
		Joint Honours	23	433	
		MSc Act. Finance	12	260	Yes
		MSc Finance	101	2,713	Yes
		MSc Civil Env Eng	13	232	
		MSc Int Health Man	43	961	
		MSc Management	118	3,756	Yes
		MSc Risk Man.	66	1,989	Yes
Pre-study Mathematics	2009	Full time MBA	(221)	1,748	
		Joint Honours	(298)	2,048	
		MSc Management	(686)	5,257	Yes
Total			1855	30,024	

4.4.3.3 *Selection of factors for the analysis hypothesis four*

Hypothesis four proposes that three factors will determine the frequency in which subjects elect to view a feedback type:

- Factor #I5 Performance
- Factor #E1 Assessment
- Factor #E5 Question focus

As described earlier, a multilevel logistic regression was used to determine whether this hypothesis was supported. Two additional factors were included in this analysis, one variable indicating the test being taken (Ethics and Strategy, Accounting Primer or Pre-study Mathematics) and one variable indicating the programme on which a subject was studying.

The individual question being presented was also initially considered for inclusion as a factor. It is rational to suppose that learners' choice to elect to view a feedback type may depend on the particular question they have answered and therefore that the individual question should be considered as a factor during the logistic regression analysis. However, given that there are 100 different questions, this would require including a categorical independent variable consisting of 100 categories. Including such a variable was not possible given the quantity of data collected. A solution to this issue would be to group the questions into a smaller number of categories. However, the two logical categories would be 'question difficulty' and 'question complexity' and both these concepts are already captured by the independent variables 'performance' and 'question focus'. The individual question was therefore not included in the logistic regression analysis as an independent variable.

4.4.3.4 Reliability and validity

Issues relating to the reliability and validity of the research design are introduced and discussed in the section 'Research methods' under 'Reliability and validity' (See page 96). The primary issue of note concerns the internal validity of the logistic regression analysis used to investigate hypothesis four. This attempts to identify a causal relationship between three factors and subjects' resultant choice to view a feedback type. As discussed, internal validity to the degree obtained via a true experimental design was not possible in this study. Neither a pre-test nor assigning subjects randomly to either a control or an experimental group were possible given that this study took place in an authentic learning context. A further consideration relating to the internal validity is the possible presence of confounding factors.

Over the course of this research, 13 factors were identified which potentially affect learners' interactions with feedback. Nine factors were incorporated into the design of this study. Three are included in the logistic regression analysis: factors '#I5 Performance', '#E1 Assessment' and '#E5 Question focus'. One factor, '#E2 Feedback level', remained constant throughout. Two factors '#I4 Level of expertise' and '#I6 Gender' were found not to be significant in studies one and four. Factors '#E3 Time limit', '#E4 Location' and '#E6 Subject matter' are incorporated into the variable 'TBI' in the logistic regression analysis. However, it was not possible to measure the individual effects of these factors for the reasons stated above under 'formation of hypotheses'.

Thus four factors remain that potentially confound the results of the logistic regression analysis. Three of these are affective factors internal to the subject: factors '#I1 Learner expectations', '#I2 Response certitude' and '#I3 Interpretation of task'. A further factor '#I7 Error type' was not incorporated into the analysis. It was not possible to control for these factors for the reasons explained above in the section 'Formation of hypotheses'. The presence of these factors is noted in the later section 'Critique of the design and conduct of the research' (See page 268).

Despite these restrictions, efforts were made regarding the internal validity of the design. The items of TBI and cohorts of students were chosen to ensure that a reasonable balance of data were captured for each independent variable. In addition, the logistic regression incorporated data from as many variables as possible in order to minimize the impact of confounding factors.

4.4.4 Results

Table 26 below shows summary data relating to the three items of TBI used in this study. These data will be referred to during the reporting of results and analysis that follows.

Table 26: Summary data relating to study three

	Ethics and Strategy exercise	Accounting Primer Course	Pre-study Mathematics Course
Number of questions in activity	50	18	25
Total number of questions responded to incorrectly	2,783 (34.35%)	2,304 (17.90%)	1,112 (12.28%)
Total number responded to correctly	5,320 (66.65%)	10,564 (82.10%)	7,941 (87.72%)
Total number of questions testing recall	4,065 (50.17%)	9,295 (72.23%)	0 (0%)
Total number of questions testing application	4,038 (49.83%)	3,573 (27.77%)	9,053 (100%)
Total number of questions (Not subsequently assessed)	8,103 (100%)	4,150 (32.25%)	3,796 (41.93%)
Total number # of questions (Subsequently assessed)	0 (0%)	8,718 (67.75%)	5,257 (58.07%)

Results for each of the six hypotheses are now presented.

4.4.4.1 Hypothesis 1: Learners will elect to view feedback more frequently than they do not

Table 27 shows the mean, median and mode α values for each feedback type together with the standard deviation. To illustrate, the table shows that on average students clicked on feedback type 'KR' for 90.79 per cent of the questions they answered. The table also shows that subjects did not view feedback for 7.50 per cent of the questions they answered and therefore viewed feedback for 92.50 per cent of the questions they answered. Hypothesis one is therefore supported.

Table 27: α values for each feedback type

	No feedback	KR	KCR	IB	EX	EI1	EI2
N	1855	1855	1855	1855	1855	1855	1855
Mean	0.0750	0.9079	0.2036	0.1257	0.1304	0.0609	0.0507
Median	0.0000	1.0000	0.1670	0.0000	0.0000	0.0000	0.0000
Std. Deviation	0.2274	0.2524	0.2494	0.2274	0.2369	0.1692	0.1518

Tables 28, 29 and 30 below show the same mean, and mode median α figures for each feedback type by test. The first table, 'Ethics and Strategy exercise' refers to the same subjects and data as study one. The three tables show

that students did not click on feedback for 14.81 per cent of the questions they answered in the Ethics and Strategy’ test, 8.43 per cent of questions they answered in the Pre-study Mathematics tests and 3.16 per cent of questions they answered in the Accounting Primer test.

Table 28: Ethics and Strategy exercise - α values for each feedback type

	No feedback	KR	KCR	IB	EX	EI1	EI2
N	146	146	146	146	146	146	146
Mean	0.1481	0.8447	0.3223	0.0655	0.0889	0.0203	0.0191
Median	0.0000	1.0000	0.3200	0.0200	0.0200	0.0000	0.0000
Std. Deviation	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 29: Pre-Study Mathematics course - α values for each feedback type

	No feedback	KR	KCR	IB	EX	EI1	EI2
N	1205	1205	1205	1205	1205	1205	1205
Mean	0.0843	0.8942	0.1638	0.1199	0.1439	0.0529	0.0373
Median	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Std. Deviation	0.2411	0.2703	0.2527	0.2355	0.2581	0.1684	0.1347

Table 30: Accounting Primer course - α values for each feedback type

	No feedback	KR	KCR	IB	EX	EI1	EI2
N	504	504	504	504	504	504	504
Mean	0.0316	0.9588	0.2642	0.1570	0.1100	0.0917	0.0918
Median	0.0000	1.0000	0.2000	0.0910	0.0000	0.0000	0.0000
Std. Deviation	0.1310	0.1538	0.2276	0.2271	0.1931	0.1874	0.1948

Thus hypothesis one is supported for each individual item of TBI. However the α value for the ‘no feedback’ condition varies noticeably across the items of TBI. The α values for the six items of feedback too vary significantly and reasons for this are discussed in the analysis of hypothesis four below.

4.4.4.2 Hypothesis 2: Subjects will elect to view KCR and KR type feedback more frequently than they elect to view either EX, IB or EI type feedback.

The α value for IB type feedback is 0.1257 meaning that, on average, subjects elected to view the relevant section of the original instruction for 12.57 per cent of the questions they answered. This α value for EX type feedback is 0.1304 and for EI1 and EI2 type feedback is 0.0609 and 0.0507 respectively. The α value for all four feedback types is lower than that of both KR (0.9079) and KCR (0.2036) type feedback and hypothesis two is supported.

Hypothesis two is also supported when considered separately for each item of TBI. For each item, subjects elected to view KCR and KR type feedback more frequently than they elect to view either EX, IB or EI type feedback. However, the extent of the difference does vary by test. In particular, the α value for EX for the Pre-study Mathematics tests (0.1439) is noticeably closer to the value for KCR (0.1638). Again, this will be discussed further in the analysis of hypothesis four below.

Relationships between feedback types

Study one found that subjects rarely elected to view both IB and EX feedback types. In study one, subjects selected either IB or EX type feedback on 9.44 per cent of occasions but selected both IB and EX type feedback on just 1.97 per cent of occasions.

As shown in table 31, subjects in this study selected either IB or EX type feedback on 9.62 per cent of occasions but selected both IB and EX type feedback on 5.09 per cent of occasions. Again, subjects elected to view either IB or EX type feedback more often than the elected to view both however the difference is less pronounced.

Table 31: The interplay between IB and EX type feedback

	Frequency	%	Cumulative %
Viewed neither IB nor EX	26,004	85.29	85.29
Viewed IB only	1,561	5.12	90.41
Viewed EX only	1,373	4.50	94.91
Viewed both IB and EX	1,552	5.09	100.00
	30,490	100.00	

Study one found that subjects rarely elected to view KCR type feedback without having elected to view KR feedback first. Table 32 shows that this is also the case in this study.

Table 32: Relationship between KR and KCR type feedback

	Frequency	%	Cumulative %
Viewed neither KR nor KCR	2,790	9.15	9.15
Viewed KR only	20,874	68.46	77.61
Viewed KCR only	168	0.55	78.16
Viewed both KR and KCR	6,658	21.84	100.00
	30,490	100.00	

4.4.4.3 *Hypothesis 2.1: For elaborative feedback types beyond KCR: Subjects will elect to view EX most frequently followed by IB and then EI type feedback.*

Study one found that subjects viewed EX most frequently followed by IB, EI1 and then EI2 type feedback. Table 33 shows that this is also the case for the Pre-study Mathematics course but not for the Accounting Primer course. Thus this hypothesis is not supported for all items of TBI.

Table 33: α values for IB, EX, EI1 and EI2 type feedback

	IB	EX	EI1	EI2
Ethics and Strategy	0.0655	0.0889	0.0203	0.0191
Pre-study Mathematics course	0.1199	0.1439	0.0529	0.0373
Accounting Primer course	0.1570	0.1100	0.0917	0.0918
All	0.1257	0.1304	0.0609	0.0507

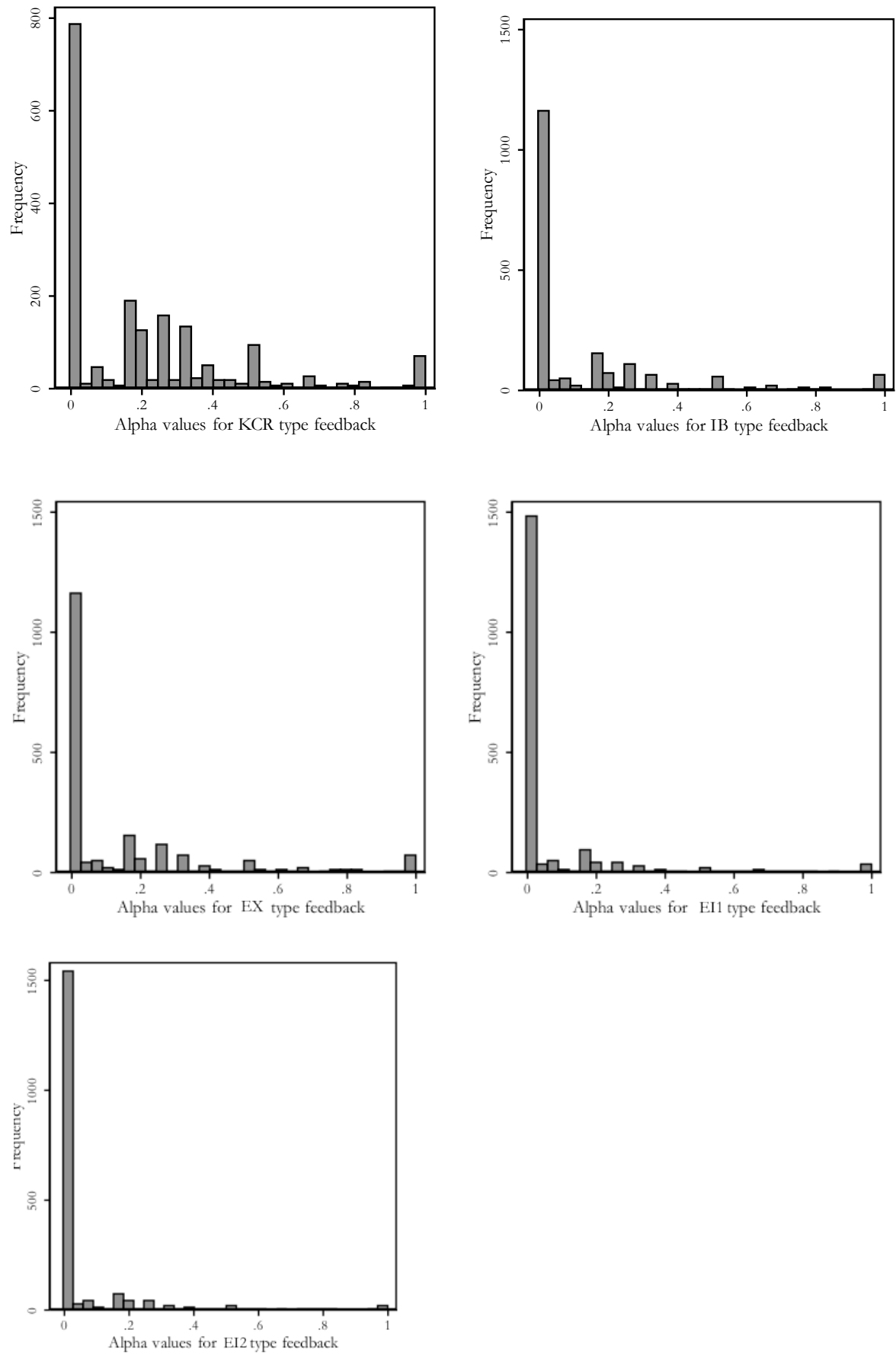
Study one also found that subjects elected to view IB and EX type feedback noticeably more frequently than they elected to view either EI1 or EI2 type feedback. Table 33 shows that this is also the case in this study for all three items of TBI, however, the difference is noticeably smaller for the Accounting Primer course.

4.4.4.4 *Hypothesis 2.2: The distribution of α values for elaborative feedback types will contain a marked positive skew*

When investigating the distribution of α values across the subject cohort, study one found that the distributions contained a marked positive skew for the KCR, IB, EX, EI1 and EI2 elaborative feedback types. A majority of subjects viewed these feedback types rarely and a minority viewed these feedback types at a wider range of frequencies.

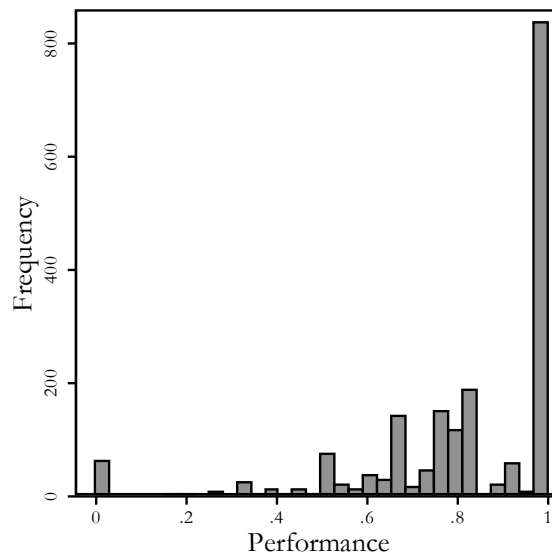
Histograms showing the frequency distributions of α values for feedback types IB, EX, EI1 and EI2 for this study are shown in figure 45. These histograms plot number of students on the y-axis and α value on the x-axis. As in study one, the distributions have a marked positive skew.

Figure 45: The distribution of α values for KCR, IB, EX, EI1 and EI2 type feedback



These histograms closely resemble the mirror image of that for performance as shown in figure 46 below. In this study, Factor #15 Performance accounts for much of the variation in the α values across the cohort. This is discussed further under ‘hypothesis four’ below.

Figure 46: Histogram showing the distribution of performance scores (%) across subjects



A significant number of subjects did not elect to view these feedback types during the activity. This is explored further in the tables below that consider just those subjects who could be identified. When anonymous submissions relating to the Pre-study Mathematics course are removed, 650 records remain. For these students it is possible to identify those who did not elect to view a particular feedback type throughout the activities. As can be seen in table 34 below, a significant proportion of subjects did not view feedback types IB, EX, EI1 and EI2 at all. Given that just 103 subjects achieved a perfect score, a significant number of subjects did not view these feedback types despite answering questions incorrectly.

Table 34: Number of subjects who never viewed a feedback type

	KR	KCR	IB	EX	EI1	EI2
Never viewed	13	93	279	346	399	422
	2.00%	14.31%	42.92%	53.23%	61.38%	64.92%
Viewed	637	557	371	304	251	228
	98.00%	85.69%	57.08%	46.77%	38.62%	35.08%
Total students	650					

Tables 35 and 36 present these same results by test.

Table 35: Ethics and Strategy exercise only

	KR	KCR	IB	EX	EI1	EI2
Never viewed	7	13	56	67	87	95
	4.79%	8.90%	38.36%	45.89%	59.59%	65.07%
Viewed	139	133	90	79	59	51
	95.21%	91.10%	61.64%	54.11%	40.41%	34.93%
Total students	146					

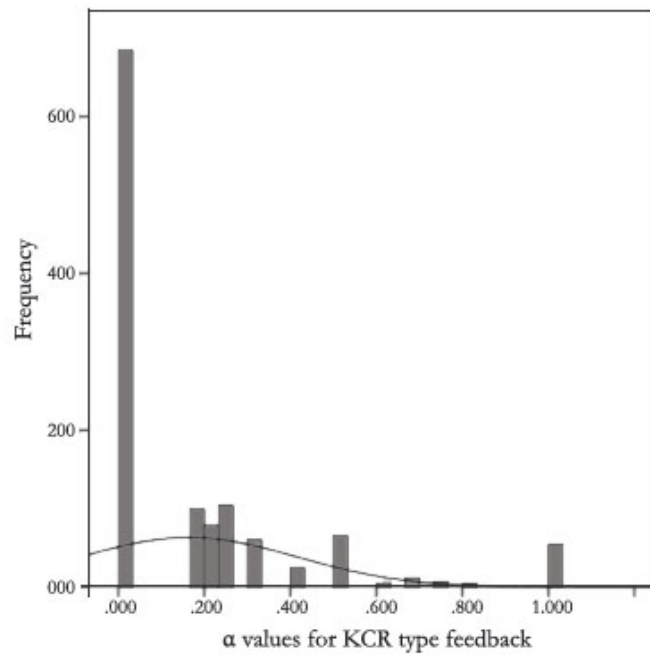
Table 36: Accounting Primer course only

	KR	KCR	IB	EX	EI1	EI2
Never viewed	6	80	223	279	312	327
	1.19%	15.87%	44.25%	55.36%	61.90%	64.88%
Viewed	498	424	281	225	192	177
	98.81%	84.13%	55.75%	44.64%	38.10%	35.12%
Total students	504					

Note: Mean and median α values for KCR type feedback in the Pre-study Mathematics tests

It is also of note that the difference between the mean (0.1638) and median (0.000) α values for KCR type feedback for the Pre-study Mathematics course is noticeably more pronounced than that for the Ethics and Strategy exercise and the Accounting Primer course. This is explained by the fact that only 12 per cent of questions were answered incorrectly in the Pre-study Mathematics tests as opposed to 18 per cent for the Accounting Primer test and 34 per cent for the Ethics and Strategy test. This moves the mode closer to the median of 0 and gives the KCR feedback type a frequency profile similar to that of the IB, EX, EI1 and EI2 feedback types. This is shown in figure 47.

Figure 47: The distribution of α values for KCR type feedback - Pre-study Mathematics



4.4.4.5 Hypothesis 3: Subjects will elect to view immediate rather than delayed feedback

Unfortunately an error in the code of the online delivery platform resulted in no data relating to feedback timing being captured for the subjects in this study.

4.4.4.6 Hypothesis 4: Learners' choice to view a feedback type is dependent on:

- a. Whether a question is answered correctly or incorrectly and
- b. Whether a question tests recall or application of the study material
- c. Whether the material being studied will be formally assessed

As described earlier, a multilevel logistic regression was used to determine whether this hypothesis was supported. For study three, five factors were initially included in the analysis. The three factors stated in the hypothesis together with one variable indicating the item of TBI being taken (Ethics and Strategy, Accounting Primer or Pre-study Mathematics) and one variable indicating the programme on which a subject was studying.

First, variance inflation factors (VIFs) were calculated to identify the presence of collinearity between the five independent predictor variables. The result indicated a high degree of collinearity between test and programme as shown in table 37.

Table 37: Initial variance inflation factors

Variable	VIF	1/VIF
TBI - Accounting	160.50	0.006231
TBI - Maths	134.86	0.00741
Programme 2	1.53	0.654261
Programme 3	6.43	0.155597
Programme 4	2.82	0.354562
Programme 5	5.80	0.172480
Programme 6	6.89	0.145156
Programme 7	57.42	0.017
Programme 8	2.77	0.361484
Programme 9	174.19	0.0057
Programme 10	43.5	0.022
Programme 11	1.25	0.798733
	160.00	0.006250
Performance	1.06	0.944363
Question focus	1.62	0.617374
Mean VIF	50.72	

Programme was removed as a factor in the analysis resulting in acceptable values of the variance inflation factors as shown in table 38.

Table 38: Variance inflation factors following removal of 'programme' as a factor

Variable	VIF	1/VIF
TBI - Accounting	2.30	0.434125
TBI - Maths	2.18	0.459693
Question focus	1.62	0.617680
Performance	1.44	0.695092
Assessed	1.06	0.945472
Mean VIF	1.00	

Next univariate and multivariate multilevel logistic regression analyses were performed for each feedback type. The results are shown in table 39 below.

Table 39: Results of univariate and multivariate multilevel logistic regression analyses

Feedback type	Variable	Category	Univariable		Multivariable	
			Odds Ratio (95% CI)	P-value	Odds Ratio (95% CI)	P-value
NF	Performance	Incorrect	1		1	
		Correct	3.57 (2.77, 4.60)	<0.001	3.56 (2.76, 4.60)	<0.001
	Assessed	No	1		1	
		Yes	1.14 (0.81, 1.59)	0.455	1.55 (1.01, 2.25)	0.018
	Question focus	Recall	1		1	
		Application	1.01 (0.83, 1.23)	0.915	0.78 (0.62, 0.99)	0.038
	TBI	Ethics	1		1	
		Accounting	0.20 (0.10, 0.40)		0.11 (0.05, 0.24)	
KR	Performance	Incorrect	1		1	
		Correct	0.61 (0.49, 0.76)	<0.001	0.62 (0.50, 0.77)	<0.001
	Assessed	No	1		1	
		Yes	0.68 (0.49, 0.94)	0.019	0.58 (0.40, 0.84)	0.004
	Question focus	Recall	1		1	
		Application	0.87 (0.71, 1.06)	0.175	1.10 (0.88, 1.38)	0.390
	TBI	Ethics	1		1	
		Accounting	1.90 (0.94, 3.86)		4.29 (2.01, 9.18)	
KCR	Performance	Incorrect	1		1	
		Correct	0.001 (0.001, 0.002)	<0.001	0.001 (0.001, 0.002)	<0.001
	Assessed	No	1		1	
		Yes	0.73 (0.67, 0.82)	<0.001	0.73 (0.60, 0.88)	0.001
	Question focus	Recall	1		1	
		Application	1.20 (1.12, 1.29)	<0.001	1.57 (1.37, 1.79)	<0.001
	TBI	Ethics	1		1	
		Accounting	0.68 (0.55, 0.82)		3.10 (2.11, 4.57)	
IB	Performance	Incorrect	1		1	
		Correct	0.11 (0.99, 0.13)	<0.001	0.11 (0.99, 0.13)	<0.001
	Assessed	No	1		1	
		Yes	1.14 (0.97, 1.33)	0.110	1.03 (0.86, 1.23)	0.745
	Question focus	Recall	1		1	
		Application	1.51 (1.35, 1.68)	<0.001	1.64 (1.44, 1.86)	<0.001
	TBI	Ethics	1		1	

		Accounting	3.12 (2.24, 4.37)		6.05 (4.12, 8.88)	
		Maths	2.37 (1.69, 3.34)	<0.001	3.78 (2.56, 5.57)	<0.001
EX	Performance	Incorrect	1		1	
		Correct	0.08 (0.07, 0.09)	<0.001	0.07 (0.07, 0.09)	<0.001
	Assessed	No	1		1	
		Yes	0.86 (0.73, 1.01)	0.072	0.86 (0.71, 1.03)	0.102
	Question focus	Recall	1		1	
		Application	2.53 (2.26, 2.86)	<0.001	2.28 (1.98, 2.64)	<0.001
	TBI	Ethics	1		1	
		Accounting	1.73 (1.22, 2.45)		4.07 (2.71, 6.10)	
		Maths	3.25 (2.29, 4.60)	<0.001	6.06 (4.04, 9.09)	<0.001
EI1	Performance	Incorrect	1		1	
		Correct	0.24 (0.20, 0.28)	<0.001	0.24 (0.21, 0.29)	<0.001
	Assessed	No	1		1	
		Yes	1.30 (1.04, 1.61)	0.017	1.08 (0.86, 1.35)	0.521
	Question focus	Recall	1		1	
		Application	1.34 (1.14, 1.57)	<0.001	1.69 (1.41, 2.02)	<0.001
	TBI	Ethics	1		1	
		Accounting	4.62 (2.96, 7.22)		6.85 (4.18, 11.22)	
		Maths	2.49 (1.57, 3.95)	<0.001	2.79 (1.68, 4.64)	<0.001
EI2	Performance	Incorrect	1		1	
		Correct	0.30 (0.25, 0.36)	<0.001	0.31 (0.26, 0.37)	<0.001
	Assessed	No	1		1	
		Yes	0.95 (0.75, 1.20)	0.669	0.72 (0.57, 0.92)	0.008
	Question focus	Recall	1		1	
		Application	1.06 (0.89, 1.25)	0.522	1.48 (1.22, 1.78)	<0.001
	TBI	Ethics	1		1	
		Accounting	5.77 (3.55, 9.38)		10.19 (6.02, 17.24)	
		Maths	2.19 (1.31, 3.62)	<0.001	2.98 (1.73, 5.14)	<0.001

The odds ratios for factors significant to a p level of less than 0.05 are summarized in the table 40 below.

Table 40: Summary of results for factors significant to a p level of less than 0.05

	TBI									
	Performance		Assessed		Ques. focus		Ethics	Acc	Maths	
	OR	p	OR	p	OR	p		OR	OR	p
NF	3.56	<0.001	1.55	0.018	0.78	0.038	1	0.11	0.58	<0.001
KR	0.62	<0.001	0.58	0.004			1	4.29	0.72	<0.001
KCR	0.001	<0.001	0.73	<0.001	1.57	<0.001	1	3.1	1.12	<0.001
IB	0.11	<0.001			1.64	<0.001	1	6.05	3.78	<0.001
EX	0.07	<0.001			2.28	<0.001	1	4.07	6.06	<0.001
EI1	0.24	<0.001			1.69	<0.001	1	6.85	2.79	<0.001
EI2	0.31	<0.001	0.72	0.008	1.48	<0.001	1	10.19	2.98	<0.001

The most significant effect was that of the variable, ‘TBI’. The odds ratios across the feedback types ranged from 0.11 to 10.19. Subjects were far more likely to view feedback types IB, EX, EI1 and EI2 when studying the Accounting Primer and Pre-study Mathematics courses than when studying the Ethics and Strategy exercise.

The results show that the variable ‘performance’ had a significant effect on subjects electing to view each one of the feedback types. Having answered a question incorrectly subjects were more likely to view feedback type KR and much more likely to view feedback types KCR, IB, EX, EI1 and EI2. The corollary being that after answering a question correctly, subjects were much more likely not to view feedback.

Whether the question being answered tests application or recall of the material studied had a significant effect on whether subjects elected to view feedback types KCR, IB, EX, EI1 and EI2. Subjects were more likely to view these feedback types if a question tested application of the material studied. The odds ratio ranged from 1.48 for EI2 type feedback to 2.28 for EX and averaged 1.73 across the feedback types. In other words the odds of a subject electing to view one of these feedback types was, on average, 73 per cent higher for questions that tested application of the material studied. In addition, subjects were less likely not to view any of the feedback types for such questions.

Whether the material being studied is subsequently assessed has a significant effect on whether subjects elected to view feedback types KR, KCR and EI2 or chose not to view feedback. While the effect was not as strong as that for the variable ‘performance’, subjects were less likely to view feedback types KR, KCR and EI2 if the material was subsequently assessed and therefore more likely not to view any feedback.

4.4.4.7 Hypothesis 5: The majority of students will revisit a previously viewed question.

As described in study one (See page 139), the online system tracks subjects’ pathways through the test and an analysis of these pathways reveals whether subjects chose to revisit a question. In study one, 86.30 per cent of

students revisited at least one question over the course of the activity and on average, subjects revisited 4.17 questions. In addition, subjects elected to view an additional feedback item on 42.19 per cent of revisits.

Unlike the Ethics and Strategy exercise, the multiple choice questions in the Accounting Primer and Pre-study Mathematics courses are divided into sections. In study three, the pathways of subjects from one section of the Accounting Primer course (The balance sheet) and one section of the Pre-Study Mathematics course (Basic Maths 2) were selected for analysis. Both sections contained tests consisting of six questions.

The following cohorts of subjects who studied the balance sheet component of the Accounting Primer course were selected for analysis.

Table 41: Subjects selected for analysis by cohort

	Frequency
FTMBA	44
IBSc	36
MSc Finance	122
Total	202

For the maths course a random sample of 50 of the 493 submissions was selected for analysis. These were obtained by placing the records in date order and then selecting every tenth submission starting from the fourth submission.

As shown in table 42 below, on average 42.85 per cent of subjects revisited at least one question and 19.84 per cent of subjects both revisited at least one question and viewed an additional item of feedback when they did so. Across the whole cohort subjects revisited 1.42 questions on average and 0.38 items of additional feedback.

Table 42: Subjects revisiting questions

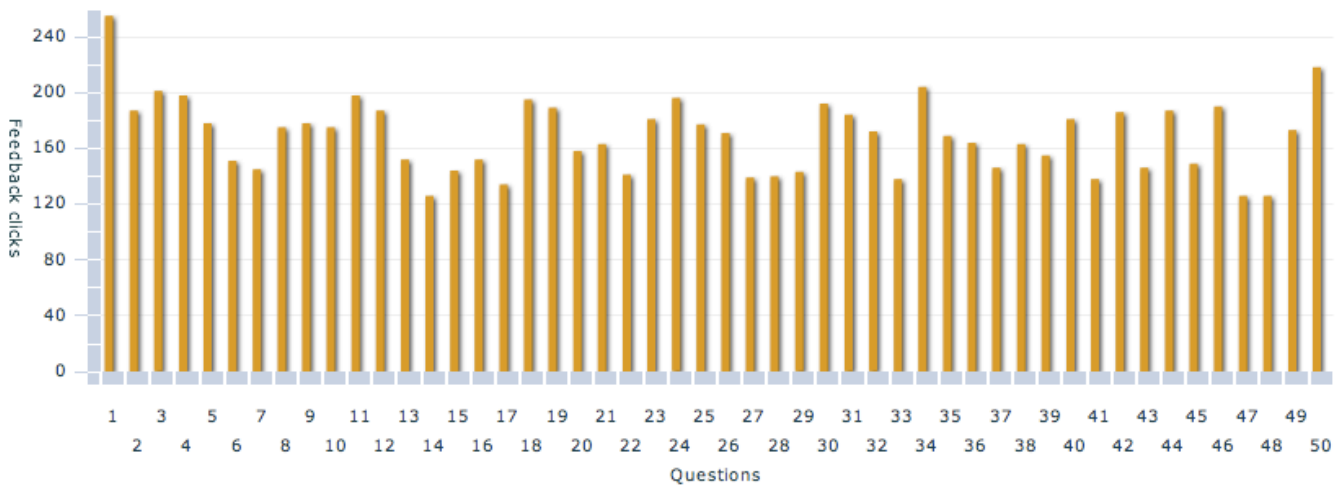
	Accounting Primer course section	Pre-study Mathematics course section	Mean across the two sections
Proportion of students who revisited at least one question	0.4356	0.4000	0.4285
Proportion of students who revisited question and viewed additional feedback	0.2228	0.1000	0.1984
Mean number of questions revisited across the cohort	2.10	0.74	1.83
Mean number of additional feedback items viewed across the cohort	0.56	0.20	0.49

In this case hypothesis five is not supported. The majority of subjects did not revisit a question.

4.4.4.8 Hypothesis 6: Learners use of feedback will not change significantly over the course of the activity

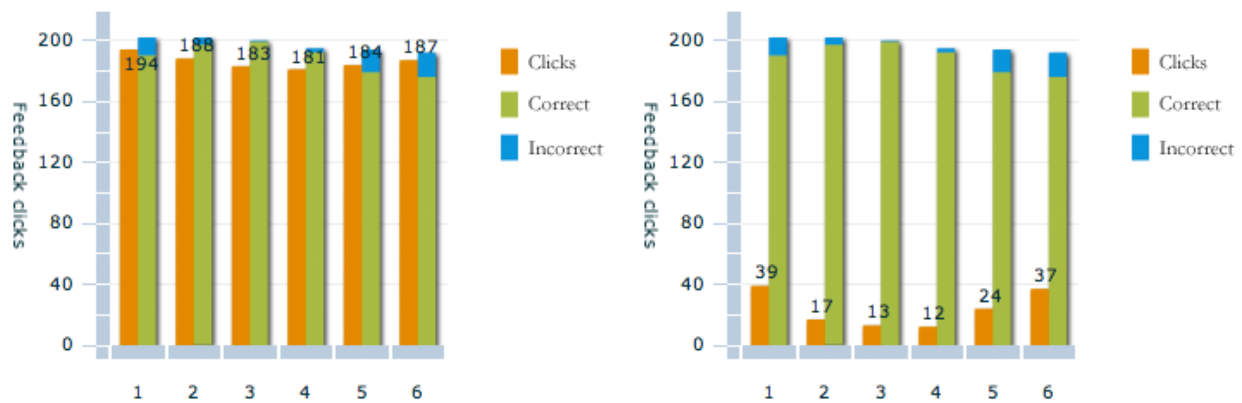
In study three, diagrams showing subjects' use of the feedback types on each of the 50 questions in the 'Ethics and Strategy' test indicated that subjects use of the feedback types did not diminish or increase over the course of the activity but remained stable. Figure 48 below shows an example of these diagrams.

Figure 48: Feedback clicks for KR type feedback across the 'Ethics & Strategy' test



The Pre-study Mathematics tests and the Accounting Primer tests are divided into smaller tests comprising between five and six questions. These smaller tests are spread throughout the TBI and thus students only sit between five and six questions at a time. Examples of the diagrams for these smaller tests are shown in figure 49.

Figure 49: Feedback clicks for KR & KCR type feedback



As these diagrams demonstrate, there are insufficient questions in these smaller tests to reveal any patterns over the course of the test. Hypothesis 6 is neither supported nor not supported.

4.4.5 Analysis and conclusions

4.4.5.1 *Subjects viewed feedback extensively.*

It is apparent from studies one and three that, if provided, students will view feedback extensively. In study one, subjects elected to view at least one feedback type for 85.19 per cent of the questions they answered. In study three, this figure rose to 92.50 per cent with individual figures for the Accounting Primer and Pre-study Mathematics courses being 96.84 per cent and 91.57 per cent respectively.

The proportion of questions answered correctly was noticeably lower for the Ethics and Strategy exercise (66.65%) than for the Accounting Primer (82.10%) and Pre-study Mathematics (87.72%) courses. Given that the analysis of hypothesis four found that subjects were more likely to view all feedback types following an incorrect response then we may expect subjects to view feedback more often during the Ethics and Strategy exercise however the opposite occurred. A possible explanation is the presence of a time limit. Subjects studying the Ethics and Strategy exercise' did so together in a computer room. Despite being given significantly more time than required, the subjects from study one who attended interviews in study two made frequent reference to the presence of a time limit being a factor affecting their choice to view the feedback types. Subjects studying the Accounting Primer and Pre-study Mathematics courses did so at a time and place of their choosing and without a time limit. This suggests that such exercises may best be conducted outside of the classroom.

4.4.5.2 *The existence of a threshold*

As described in the analysis of study one, prior research has proposed the existence of a threshold, at the level of KCR type elaborative feedback, beyond which the study of feedback becomes either ineffective or inefficient (Phye 1979; Kulhavy, White et al. 1985; Dempsey and Litchfield 1993; Pridemore and Klein 1995). This proposal was supported by the results of study one which found that subjects elected to view feedback types KR and KCR noticeably more frequently than types IB and EX which in turn were viewed more frequently than types EI1 and EI2.

In study three, subjects again viewed feedback types KR and KCR noticeably more frequently than the other feedback types however the difference was not as striking and does not give the appearance of a threshold. Study three included data from two additional items of TBI, the Accounting Primer course and the Pre-study Mathematics course. For both items, subjects elected to view KCR feedback noticeably more frequently than the other elaborative feedback types however the difference when compared to the IB and EX feedback types was not as pronounced as in study one. In particular, the difference in α values between KCR type feedback and EX type feedback for subjects taking the Pre-study Mathematics course is noticeably smaller. These results are summarised in table 43 below.

Table 43: Mean α values for each feedback type by item of TBI

	% correct responses	No feedback	KR	KCR	IB	EX	EI1	EI2
Ethics and Strategy exercise	66.65%	0.1481	0.8447	0.3223	0.0655	0.0889	0.0203	0.0191
Pre-Study Mathematics course	87.72%	0.0843	0.8942	0.1638	0.1199	0.1439	0.0529	0.0373
Accounting Primer Course	82.10%	0.0316	0.9588	0.2642	0.1570	0.1100	0.0917	0.0918

Table 44 below shows the same information but expressed in terms of number of questions.

Table 44: Mean α values expressed as number of questions

	No feedback	KR	KCR	IB	EX	EI1	EI2
Ethics and Strategy exercise	7.40	42.23	16.12	3.27	4.44	1.01	0.95
Pre-Study Mathematics course	2.11	22.36	4.10	3.00	3.60	1.32	0.93
Accounting Primer course	0.57	17.26	4.76	2.83	1.98	1.65	1.65

Subjects were far more likely to view feedback types IB, EX, EI1 and EI2 when studying the Accounting Primer and Pre-study Mathematics courses than when studying the Ethics and Strategy exercise. The former courses contained a higher proportion of questions testing application rather than recall of the course material and the difference is partly explained by the results of hypothesis four below in which subjects were more likely to view feedback types KCR, IB, EX, EI1 and EI2 if a question tested application of the material studied.

4.4.5.3 *Subjects held a preference for feedback that focuses on the knowledge gap*

A hypothesis emerged in study one that subjects held a preference for elaborative feedback beyond KCR that focuses specifically on the knowledge gap, that is the gap between what was understood and what had been required to be understood in order to have answered a question correctly. Study two provided some support for this hypothesis. Should this hypothesis hold in study three then we should expect mean α values for KCR type feedback to be the highest followed by those for EX then IB and then the extra-instructional feedback types EI1 and EI2. As can be seen in table 45 below, when mean α values for all items of TBI combined are considered, this hypothesis is supported.

Table 45: Mean α values for each feedback type – All items of TBI

	No feedback	KR	KCR	IB	EX	EI1	EI2
Mean	.07496	.90787	.20358	.12570	.13035	.06089	.05070

However when mean α values are considered separately for each test, the hypothesised order does not hold for the Accounting Primer Course. This is shown in table 46.

Table 46: Mean α values for each feedback type – For each item of TBI

	No feedback	KR	KCR	IB	EX	EI1	EI2
Ethics and Strategy exercise	0.1481	0.8447	0.3223	0.0655	0.0889	0.0203	0.0191
Pre-Study Mathematics course	0.0843	0.8942	0.1638	0.1199	0.1439	0.0529	0.0373
Accounting Primer Course	0.0316	0.9588	0.2642	0.1570	0.1100	0.0917	0.0918

The mean α values for the Accounting Primer Course differ from those predicted by the hypothesis in that the mean α value for IB type feedback is higher than that for EX type feedback

However, upon closer examination, a possible explanation is apparent. As discussed in the analysis and conclusions sections in studies one and two (See pages 144, 174), EX type feedback can be considered to be more specifically focused on the knowledge gap than IB type feedback and it was hypothesised that subjects would elect to view the former more frequently. The reasoning is as follows; by providing an explanation as to why a user's response was incorrect, EX type feedback focuses specifically on the task posed by a question. IB type feedback presents a subject with the relevant section from the course material. This feedback type is sufficient to close the knowledge gap when a question recycles concepts verbatim from the course material. However, this feedback type may not be sufficient to close the knowledge gap when a subject is required to apply concepts from the materials to the question. This feedback type can be considered to be generally less focused on the knowledge gap than EX type feedback.

However, a closer inspection of the questions from the Accounting Primer course revealed that many questions do recycle concepts verbatim from the course material. 70.82 per cent of the questions responded to by the subjects in the Accounting Primer course tested recall of the course material and, in these cases, the presentation of the original material directly answers the question. It does not require the learner to apply concepts from the materials to a novel situation. This can be seen in figure 50, a screenshot from the test that shows one such example:

Figure 50: Screenshot from the Accounting Primer course

Which of the following is not a "current liability"?

- ☒ Long term loan
- ☐ Creditors
- ☐ Provision for Tax
- ☐ Accrued expenses

Submit answer

► Did I answer correctly?

► What was the correct answer?

▼ Course reference:

Marketable securities	250			
Cash	3,550			
Prepaid expenses	400			
Total Current Assets		22,800		
Less Current Liabilities:				
Creditors	(6,800)			
Provision for Tax	(1,700)			
Accrued expenses	(600)			
Total Current Liabilities		(9,100)		
Net Current Assets				13,700
Total Assets less Current				27,200

Thus, in the case of the Accounting Primer course, IB type feedback fulfils the explanatory role as efficiently as EX type feedback for the majority of questions. In addition, by presenting the original material, the learner can readily identify the nature of any misunderstandings that arose during their study of the material. Thus, for this item of TBI, IB type feedback can generally be considered to focus more directly on any knowledge gaps than EX type feedback.

Thus while the results for Accounting Primer course do differ from those predicted, a possible explanation exists as to why this is the case. This research concludes that study three provides support for the hypothesis the subjects hold a preference for feedback that focuses specifically on the knowledge gap. This will be investigated further in study four.

4.4.5.4 Patterns in subjects' behaviour

When investigating the distribution of α values across the subject cohort, study one found a common pattern. A large majority of subjects elected to view feedback type KR for virtually all of the questions. For the elaborative feedback types, a large majority of subjects viewed these feedback types rarely while a minority viewed these feedback types at a broad range of frequencies. Thus, when plotted on a graph, the distributions of α values across the cohort contained a marked positive skew for the KCR, IB, EX, EI1 and EI2 elaborative feedback types. For KCR type feedback this difference in behaviour across the cohort was largely explained by the differences in performance. However, for feedback types IB, EX, EI1 and EI2, performance was not able to totally explain the variety in subjects' behaviour.

Study two suggested that factor ‘#I1 learner expectations’ (through subjects’ style of engagement and sense of enquiry), factor ‘#I3 interpretation of task’ and factor ‘#I4 level of expertise’ may account for the variation in the frequency at which subjects viewed the elaborative feedback types not explained by differences in performance.

In study three, subjects’ patterns of behaviour resembled those observed in study one. Nearly all subjects elected to view feedback type KR for virtually all of the questions. For the elaborative feedback types, the distribution of α values across the cohort again contained a marked positive skew. A large majority of subjects viewed these feedback types rarely while a minority viewed these feedback types across a wide range of frequencies. As in study one, factor ‘#I5 performance’ largely determined this variation but not completely.

4.4.5.5 Factors affecting subjects’ choice to view a feedback type

The effect of Factor #I5 Performance

Study one found that performance had a significant and sizable effect on whether a subject elected to view a feedback types KCR, IB and EX. The effect of performance was more pronounced in study three. Having answered a question incorrectly subjects were more likely to view feedback type KR and much more likely to view feedback types KCR, IB, EX, EI1 and EI2. Conversely, after answering a question correctly, subjects were much more likely not to view feedback.

The effect of Factor #E5 Question focus

The interviews in study two suggested that the focus of a question, whether a question tests recall or application or the course material, had affected subjects’ choice to view a feedback type in study one. This factor, termed ‘#E5 question focus’, was incorporated into study three and its effect was found to be significant. Subjects were significantly more likely to view feedback types KCR, IB, EX, EI1 and EI2 if a question tested application of the material studied. A possible explanation for this is that application of a concept requires novel procedures not presented explicitly in the study material. For example, determining the correct response to quantitative questions often involves a process, a sequence of steps, which differ in detail to those presented in the material. KCR type feedback alone does not reveal these novel procedures.

The effect of Factor #E1 Assessment

Factor #E1 assessment was also found to be a significant factor. If the course material was subsequently assessed, then subjects were significantly less likely to view feedback types KR (OR=0.58), KCR (OR=0.73) and EI2 (OR=0.72) and consequently more likely to not to view feedback at all (OR=1.55). This contradicts the finding of Morrison, Ross et al. (1995) who found that subjects were 3.24 times more likely to view a review screen containing feedback following an incorrect response when a test was assessed. A possible explanation is that the construct of the term ‘assessed’ differs across the two studies. In study three, the material being tested is subsequently assessed while in Morrison, Ross et al. (1995) the test itself is assessed. Hypothesis four supposed that a test being subsequently assessed would result in the feedback types being viewed more frequently. This was not the case and an opposite effect was found for three of the feedback types. This is a result noted as an area for further investigation in the section ‘Conclusions: Suggestions for future research’ (See page 272).

The effect of the TBI material

The largest effects in this analysis related to the TBI material. Subjects were far more likely to view the elaborative feedback types IB, EX, EI1 and EI2 when studying the Accounting Primer and Pre-study Mathematics courses than when studying the Ethics and Strategy exercise. This is not due to the differing levels or difficulty, nor the proportion of application questions, as these factors are controlled for in the logistic regression analysis.

A possible explanation for this difference was the presence of a time limit when subjects undertook the Ethics and Strategy exercise. In the interviews that followed, subjects stated the presence of a time limit to be a factor when deciding the view the feedback types. This is discussed further in the 'Discussion' chapter of this thesis (See page 257).

4.4.5.6 Subjects revisit questions and feedback

In study one, 86.3 per cent of students revisited at least one question over the course of the activity. This led to the hypothesis that the majority of students will revisit a previously viewed question. This hypothesis was not supported in study three. However this is explained by the difference in the length of the tests. The test in the Ethics and Strategy exercise from study one consists of 50 questions while the tests in the sections of the Accounting Primer and Pre-study Mathematics courses consists of up to six questions. Given the relatively short length of these latter two tests, subjects were more likely to hold memory of prior questions throughout the test. However, 42 per cent of subjects did choose to revisit a question and this is of interest, particularly to those involved in instructional design.

4.4.6 Summary

This study aimed to test the results of study one with additional cohorts of students and also to test hypotheses arising from the analyses of results in studies one and two.

Supporting the results of study one, subjects viewed feedback extensively and thus hypothesis one was supported. However it was notable that the α value for the 'no feedback' condition varied noticeably across the three items of TBI.

Some further support emerged for the existence of a feedback threshold, at the level of KCR type feedback, beyond which learners elect to view the elaborative feedback types less frequently. However, this study did reveal the presence of a threshold to be noticeably less marked for the Pre-study Mathematics course suggesting that the existence of a threshold may depend on the materials being studied.

The results provided continued support for the hypothesis that learners hold a preference for feedback that focuses specifically on the knowledge gap and this will be investigated further in study four.

The test of hypothesis four found the effect of factor '#I5 performance' to be noticeably more pronounced than in study one. Having answered a question incorrectly subjects were more likely to view feedback type KR and much

more likely to view feedback types KCR, IB, EX, EI1 and EI2. The effects of factors ‘#E1 Assessment’ and ‘#E5 Question focus’ were also found to be significant. Whether a question tests application or recall of the material studied had a significant effect on whether subjects elected to view feedback types KCR, IB, EX, EI1 and EI2 and subjects were less likely to view feedback types KR, KCR and EI2 if the material was subsequently assessed.

The patterns of behaviour observed in study three resemble those observed in study one. Virtually all subjects elected to view feedback type KR for virtually all of the questions. For the elaborative feedback types, the distribution of α values across the cohort again contained a marked positive skew. For the elaborative feedback types, particularly KCR, this variation is largely explained by differences in performance across the cohort. However, performance did not explain all the variation observed.

Fewer students chose to revisit questions than in study one. 42 per cent of subjects did so as opposed to 86 per cent in study one. However, a possible explanation is that this is due to the tests comprising fewer questions.

These results will be investigated further in study four.

4.5 STUDY FOUR: THE IMPACT OF FACTOR ‘#I4 LEVEL OF EXPERTISE’ (2010)

4.5.1 Introduction

This study comprises a technology based observation of three cohorts of students studying two different items of TBI in the academic year 2010 to 2011. As per study three, this study re-examines the hypotheses explored in study one and investigates additional hypotheses that arose during the analyses of studies one and two. The study took place between July and December 2010.

The format follows the same fundamental design as studies one and three. The three cohorts studied both the Accounting Primer and the Pre-study Mathematics courses described earlier. The most significant difference with study three was the inclusion of an additional factor ‘#I4 Level of expertise’ in the analysis of hypothesis four.

As before subjects were presented with six feedback options upon responding to a multiple choice question corresponding to the six feedback types in the taxonomy of TBI feedback. Subjects were also given the choice either to view feedback immediately following their response to a question or after having answered all of the questions. As before, the software tool recorded all ‘clicks’ made by subjects in the delivery interface.

The final hypotheses for study four are:

1. Learners will elect to view feedback more frequently than they do not
2. Learners will elect to view KCR and KR type feedback more frequently than they elect to view either EX, IB or EI type feedback
 - 2.1. For elaborative feedback type beyond KCR: Learners will elect to view EX most frequently followed by IB and then EI type feedback
 - 2.2. The distribution of α values for elaborative feedback types will contain a marked positive skew
3. Learners will elect to view immediate feedback more frequently than they elect to view delayed feedback
4. Learners’ choice to view a feedback type is dependent on:
 - Factor ‘#I5 Performance’ (Whether a question is answered correctly or incorrectly)
 - Factor ‘#E5 Question focus’ (Whether a question tests recall or application of the study material)
 - Factor ‘#I4 Level of expertise’ (The level of the learner)
 - Factor ‘#I6 Gender’ (Whether the subject is male or female.)

5. The majority of learners will revisit a previously viewed question
6. Learners' use of feedback will not change significantly over the course of the activity

The process of deriving the hypotheses for this study is now described.

4.5.2 Formation of hypotheses

The hypotheses in study four are essentially the same as those in study three (See page 180). The one difference being the factors included in hypothesis four. The process of selecting these factors is now described.

The literature review identified nine factors that potentially affect whether a learner elects to view a feedback type. Study two identified a further four factors resulting in a total of 13 possible factors that were considered for investigation in study four, as shown in table 47.

Table 47: The 13 possible factors for consideration for investigation in study four

Internal factors	External factors
Factor #I1 Learner expectations	Factor #E1 Assessment
Factor #I2 Response certitude	Factor #E2 Feedback level
Factors #I3 Interpretation of task	Factor #E3 Time limit
Factor #I4 Level of expertise	Factor #E4 Location
Factor #I5 Performance	Factor #E5 Question focus
Factor #I6 Gender	Factor #E6 Subject matter
Factor #I7 Error type	

Four of these factors were chosen for investigation through the following process of elimination. Factors ‘#I1 Learner expectations’, ‘#I2 response certitude’, ‘#I3 interpretation of task’, ‘#I8 error type’, ‘#E2 feedback level’, ‘#E3 time limit’ and ‘#E4 location’ and ‘#E6 Subject matter’ were eliminated from study four for the same reasons stated in the section ‘Formation of hypotheses’ in the report of study three (See page 181).

Factor ‘#E1 Assessment’ was not included as the three cohorts of students were not subsequently assessed on either the “Accounting Primer” course or the “Pre-study Mathematics” course.

As per study three, factors ‘#I5 Performance’ and ‘#E5 Question focus’ were included. Unlike study three, data for factors ‘#I4 Level of expertise’ and ‘#I6 Gender’ were available and these factors were also included in study four.

4.5.3 Design

This study adopts the technology based observation approach described in the section ‘Research methods’ (See page 85) and also in the reports of studies one (See page 128) and three (See page 183). However the factor constructs incorporated into hypothesis four, the selection of subjects and issues relating to reliability and validity do differ. These topics are now introduced and discussed.

4.5.3.1 Factor constructs

The additional factor investigated in this study is ‘#I4 Level of expertise’. Both the Accounting Primer and Pre-study mathematics courses contain a final assessment consisting of a multiple choice test that tests students’ ability in all sections of the course. In this study, subjects’ score in this final assessment was adopted as a measure of their level of expertise.

4.5.3.2 Selection of subjects

Data for the following 17 cohorts were captured by the software system in the academic year 2010 to 2011.

Table 48: Summary data relating to the cohorts in 2010-2011

Test	Programme	Number of subjects
Accounting Primer Course	EMBA	22
	Executive Education	24
	Full time MBA	69
	IBSc Programme	67
	Joint Honours	38
	MSc Finance	135
	MSc Actuarial Finance	14
	MSc Int. Health Man.	39
	MSc Management	145
	MSc Risk Management	77
Pre-study Mathematics Course	WEMBA	39
	DLMBA	1
	EMBA	19
	Full time MBA	80
	IBSc Programme	69
	Joint Honours	47
	WEMBA	30

In order to allow a direct comparison across the two items of TBI, cohorts who had studied both the Accounting Primer and Pre-study mathematics courses were selected. This decision resulted in five possible cohorts as candidates for selection the EMBA, FTMBA, IBSc Programme, Joint Honours and WEMBA cohorts.

Significant data processing was required for each subject and due to time constraints, three cohorts of students were chosen to be the subjects in this study, these being the three largest cohorts. As before records that contained unsolvable errors were removed, as were records for subjects for whom either gender or learner level was unknown. This resulted in the selection of subjects shown in table 49.

Table 49: The student cohorts forming the research subjects in study four

Test	Year	Cohort	Number of subjects	No. of questions answered
Accounting Primer course	2010	Full time MBA	38	831
		IBSc Programme	34	465
		Joint Honours	25	1,965
Pre-study Mathematics Course	2010	Full time MBA	45	1,390
		IBSc Programme	39	1,151
		Joint Honours	30	738

4.5.3.3 Reliability and validity

General issues relating to the reliability and validity of the research design are introduced and discussed in the section ‘Research methods’ under ‘Reliability and validity’ (See page 96). As per study three, the primary issue of note relates to the internal validity of the logistic regression analysis used to test of hypothesis four which attempts to identify a causal relationship between five factors and subjects resultant choice to view a feedback type. As described earlier (See page 98), internal validity to the degree obtained via a true experimental design was not possible in this study. As in study three, there are four factors that potentially confound the results of the logistic regression analysis. Three of these are affective factors internal to the subject: factors ‘#I1 Learner expectations’, ‘#I2 Response certitude’ and ‘#I3 Interpretation of task. A further factor #I7 Error type was not incorporated into the analysis. It was not possible to control for these factors for the reasons explained above in the section ‘Formation of hypotheses’ above. The presence of these factors is noted in the later section ‘Critique of the design and conduct of the research’.

Despite this, as per study three, efforts were made as regards the internal validity of the approach. As described above under ‘Selection of subjects’, the cohorts of students were chosen to ensure that a reasonable balance of data was captured for each independent variable. And, as per study three, the logistic regression incorporated data from as many variables as possible in order to identify confounding factors.

4.5.4 Results

Table 50 provides summary data relating to the three items of TBI used in this study. These data will be referred to during the reporting of results and analysis that follows.

Table 50: Summary data relating to the two items of TBI used in study four

	Accounting Primer		Pre-study Mathematics		Total
Number of questions answered incorrectly	389	19.54%	399	12.17%	788
Answered correctly	1,602	80.46%	2,880	87.83%	4,482
Number of questions testing recall	1,410	70.82%	0	0.00%	1,410
Number of questions testing application	581	29.18%	3,279	100.00%	3,860
Total number of questions	1,991		3,279		5,270

4.5.4.1 Hypothesis 1: Learners will elect to view feedback more frequently than they do not

Table 51 below shows the mean, and median α figures for each feedback type together with the standard deviation. To illustrate, the table shows that on average students clicked on feedback type 'KR' for 93.97 per cent of the questions they answered.

The table also shows that subjects did not view feedback for 5.33 per cent of the questions they answered and thus viewed feedback for 94.67 per cent of the questions they answered therefore hypothesis one is supported.

Table 51: α values for each feedback type

	No feedback	KR	KCR	IB	EX	EI1	EI2
N	211	211	211	211	211	211	211
Mean	0.0533	0.9397	0.2082	0.1099	0.1410	0.0496	0.0480
Median	0.0000	1.0000	0.1670	0.0400	0.0830	0.0000	0.0000
Std. Deviation	0.1676	0.1851	0.1831	0.1737	0.1867	0.1461	0.1336

Tables 52 and 53 below show the same mean and median α figures for each feedback type by item of TBI. The tables show that students did not click on feedback for 2.35 per cent of questions they answered in the Accounting Primer tests and for 7.87 per cent of questions they answered in the Pre-study Mathematics tests. Thus hypothesis one is also supported for each item of TBI.

Table 52: α values for each feedback type – Accounting Primer course only

	No feedback	KR	KCR	IB	EX	EI1	EI2
N	97	97	97	97	97	97	97
Mean	0.0235	0.9653	0.2420	0.1106	0.1261	0.0489	0.0521
Median	0.0000	1.0000	0.1820	0.0000	0.0590	0.0000	0.0000
Std. Deviation	0.1009	0.1415	0.2056	0.1930	0.2158	0.1473	0.1507

Table 53: α values for each feedback type – Pre-study Mathematics course only

	No feedback	KR	KCR	IB	EX	EI1	EI2
N	114	114	114	114	114	114	114
Mean	0.0787	0.9178	0.1794	0.1092	0.1536	0.0503	0.0445
Median	0.0000	1.0000	0.1600	0.0720	0.1200	0.0000	0.0000
Std. Deviation	0.2053	0.2137	0.1568	0.1562	0.1576	0.1457	0.1177

4.5.4.2 Hypothesis 2: Learners will elect to view KCR and KR type feedback more frequently than they elect to view either EX, IB or EI type feedback.

The mean α value for IB type feedback is 0.1099 meaning that, on average, subjects elected to view the relevant section of the original instruction for 10.99 per cent of the questions they answered. This mean α value for EX type feedback is 0.1410 and for EI1 and EI2 type feedback 0.0496 and 0.0480 respectively. The α value for all four feedback types is lower than that of both KR (0.9397) and KCR (0.2082) type feedback and therefore hypothesis 2 is supported.

Hypothesis 2 is also supported when looking at the data by item of TBI. For each, subjects elected to view KCR and KR type feedback more frequently than they elect to view either EX, IB or EI type feedback. However, as per study three, the value for EX for the Pre-study Mathematics course (0.1536) is noticeably closer to the value for KCR (0.1794). This is investigated further in the discussion of hypothesis four below.

Relationships between feedback types

Studies one and three found that subjects elected to view both IB and EX feedback types less frequently than they elected to view just one of these types. In study one, subjects selected either IB or EX type feedback on 9.44 per cent of occasions but selected both IB and EX type feedback on just 1.97 per cent of occasions. In study three, subjects selected either IB or EX type feedback on 9.62 per cent of occasions but selected both IB and EX type feedback on 5.09 per cent of occasions. The data for study four are presented in table 54.

Table 54: The interplay between IB and EX type feedback

	Frequency	Percentage
Viewed neither IB nor EX	4,406	83.61%
Viewed IB only	158	3.00%
Viewed EX only	360	6.83%
Viewed both IB and EX	346	6.57%
	5,270	

In this study, subjects selected either IB or EX type feedback on 9.83 per cent of occasions but selected both IB and EX type feedback on 6.57 per cent of occasions. Again, subjects elected to view either IB or EX type feedback more often than they elected to view both. However the difference is less pronounced.

As in studies one and two, subjects rarely elected to view KCR type feedback without having elected to view KR feedback first. This is shown in table 55 below.

Table 55: The interplay between IB and EX type feedback

	Frequency	Percentage
Viewed neither KR nor KCR	361	6.85%
Viewed KR only	3,924	74.46%
Viewed KCR only	17	0.32%
Viewed both KR and KCR	968	18.37%
	5,270	

4.5.4.3 Hypothesis 2.1: For elaborative feedback type beyond KCR: Learners will elect to view EX most frequently followed by IB and then EI type feedback.

Study one found that subjects viewed EX most frequently followed by IB and then EI type feedback. This led to the formation of hypothesis 2.1. Study three found the hypothesis was supported for the Ethics and Strategy exercise and the Pre-study Mathematics course but not for the Accounting Primer course. In the latter case, subjects viewed IB type feedback more frequently than EX type feedback.

In study four, as shown in Table 56 below, hypothesis 2.1 was found to hold for both the Pre-study Mathematics course and the Accounting Primer course.

Table 56: Mean α values for IB, EX, EI1 and EI2 type feedback

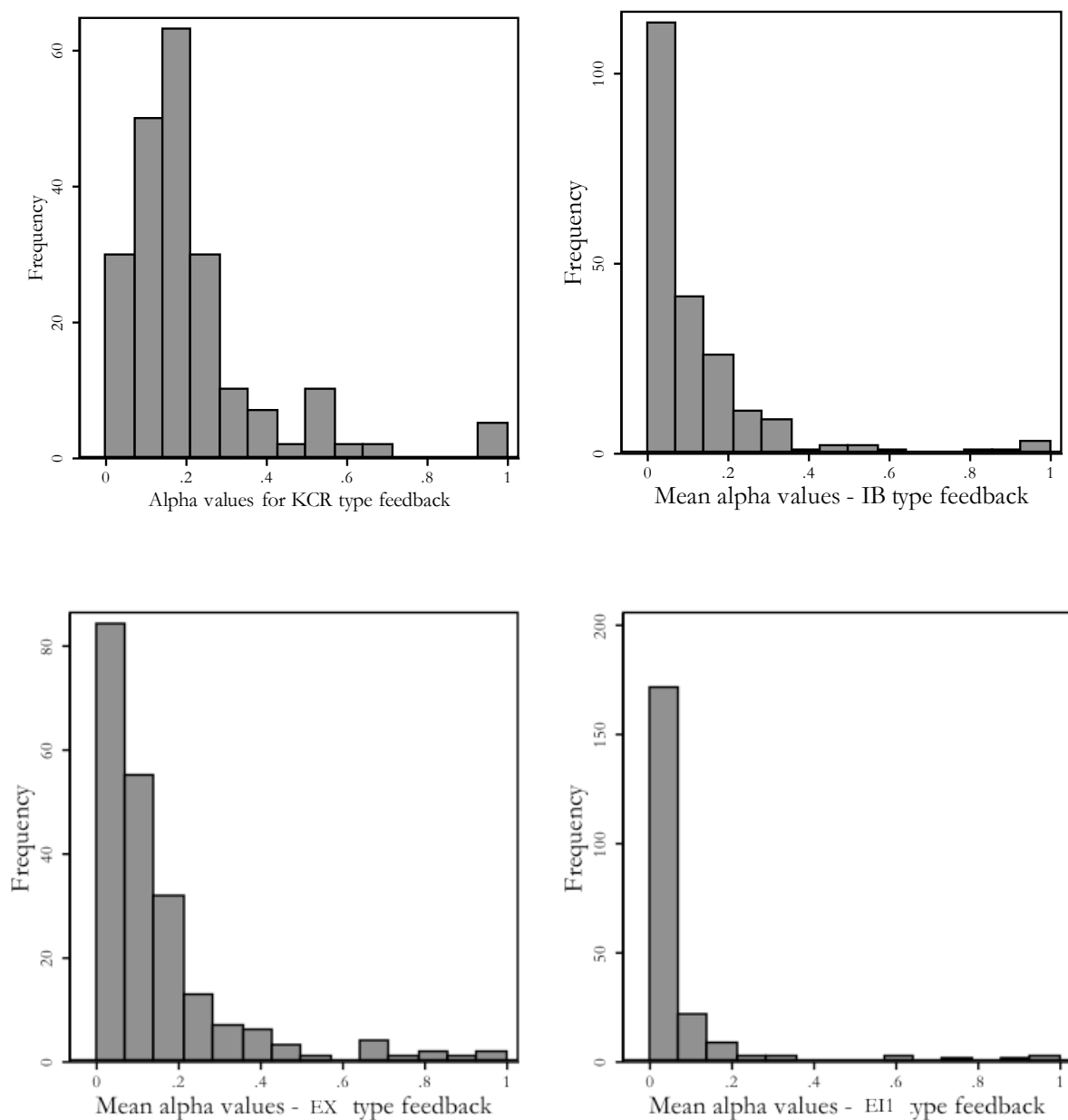
	IB	EX	EI1	EI2
All	0.1099	0.141	0.0496	0.048
Accounting Primer	0.1106	0.1261	0.0489	0.0521
Pre-study Mathematics	0.1092	0.1536	0.0503	0.0445

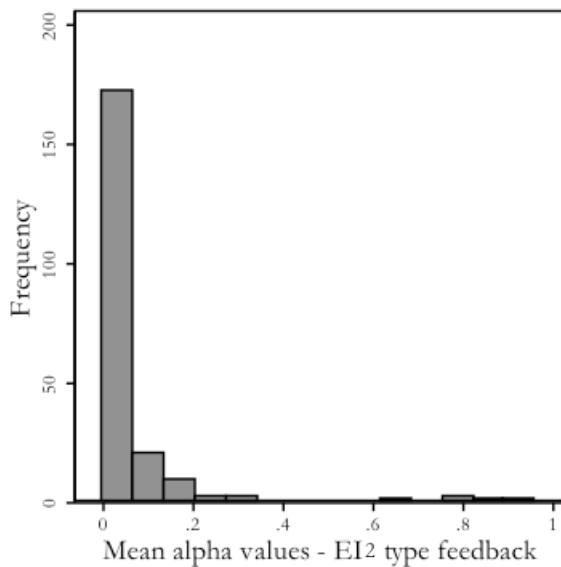
4.5.4.4 Hypothesis 2.2: The distribution of α values for elaborative feedback types will contain a marked positive skew

When investigating the distribution of α values across the subject cohort, study one found that the distributions contained a marked positive skew for the KCR, IB, EX, EI1 and EI2 elaborative feedback types.

Histograms showing the frequency distributions of mean α values for the KCR, IB, EX, EI1 and EI2 elaborative feedback types are shown in figure 51. As per studies one and three, these distributions contain a marked positive skew. A significant proportion of the subjects did not elect to view these feedback types, the majority viewed them rarely and a minority of students elected to view these feedback types across a range of frequencies.

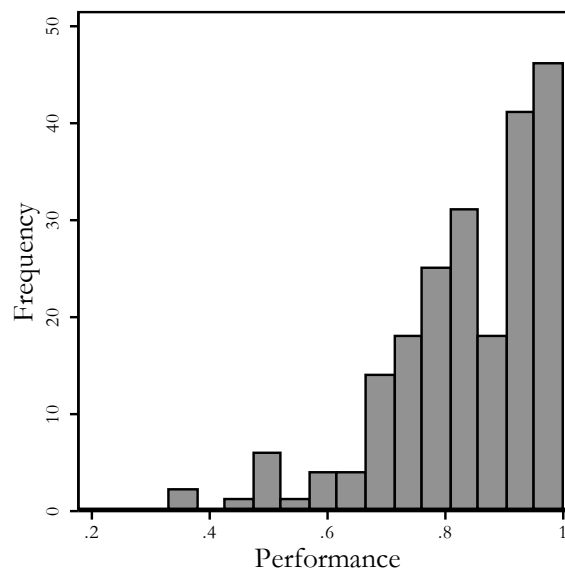
Figure 51: Distribution of mean α values for KCR, IB, EX, EI1 and EI2 feedback types





In study three, the variation in subjects' behaviour was largely explained by differences in performance across the cohort. In this study, the frequency distributions for the elaborative feedback types shown above again closely resemble the mirror image of the frequency distribution for performance scores as shown in figure 52 below. As in study three, performance accounts for much of the variation in α values across the cohort. This is discussed further in the analysis of 'hypothesis four' below.

Figure 52: The frequency distribution of performance scores (%) across all subjects



Note on the distribution of α values for feedback types IB and EX for subjects studying the Pre-study Mathematics tests

When results are considered separately for each test it is clear that median α values for feedback types IB and EX are noticeably closer to the mean for subjects studying the Pre-study Mathematics tests indicating a reduction in the positive skew in the distribution of these values.

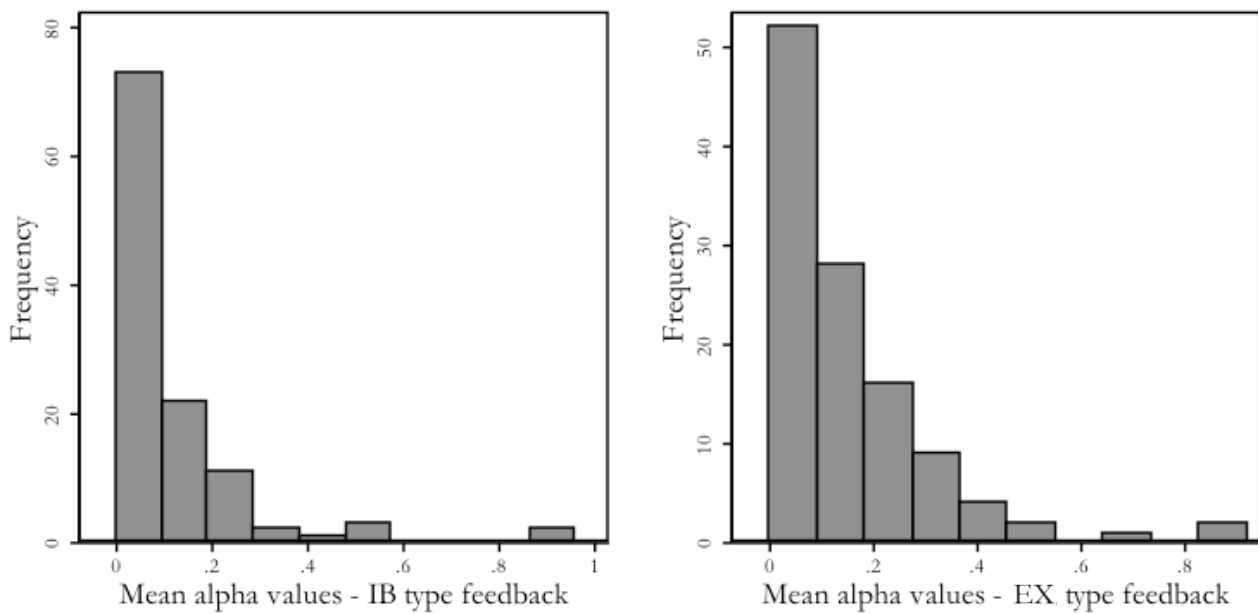
Table 57: Accounting Primer Course only

	No feedback	KR	KCR	IB	EX	EI1	EI2
Mean	0.0235	0.9653	0.2420	0.1106	0.1261	0.0489	0.0521
Median	0.0000	1.0000	0.1820	0.0000	0.0590	0.0000	0.0000

Table 58: Pre-study Mathematics only

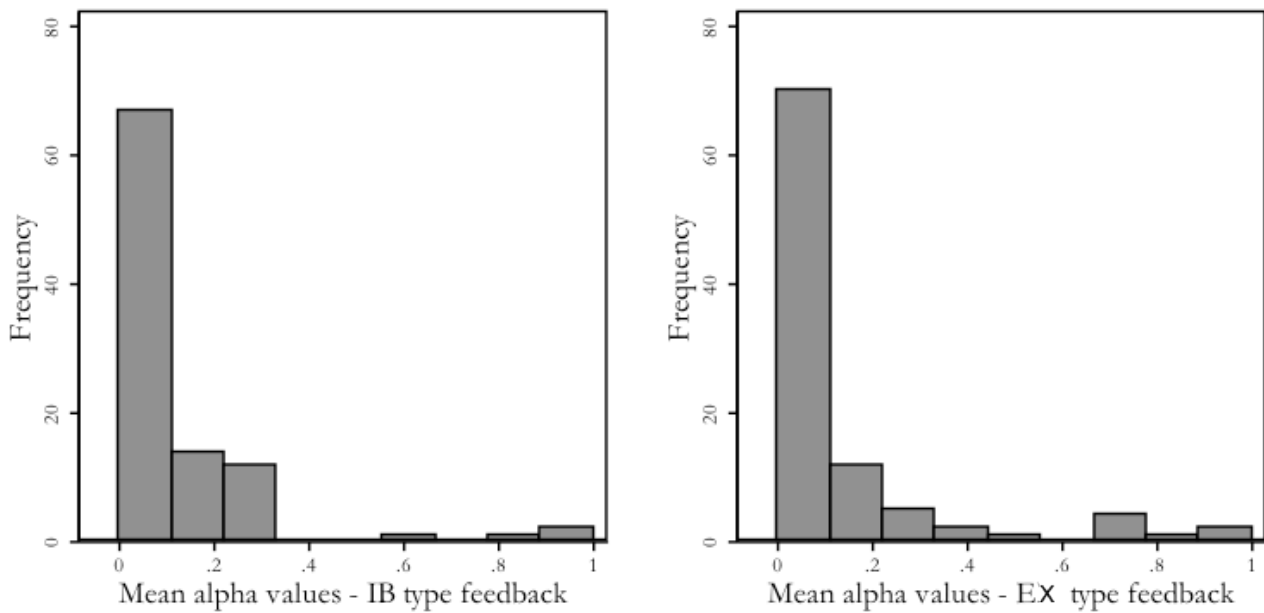
	No feedback	KR	KCR	IB	EX	EI1	EI2
Mean	0.0787	0.9178	0.1794	0.1092	0.1536	0.0503	0.0445
Median	0.0000	1.0000	0.1600	0.0720	0.1200	0.0000	0.0000

This is illustrated in figure 53 which contains histograms showing the distribution of the mean α values for feedback types IB and EX for the Pre-study Mathematics tests alone.

Figure 53: Distributions of α values for IB/EX feedback types - Pre-study Mathematics

As shown in figure 54, the distribution of the mean α values for feedback types, IB and EX for the Accounting Primer tests is similar to the distributions seen in studies one and two.

Figure 54: Distributions of α values for IB/EX feedback types - Accounting Primer course



4.5.4.5 Hypothesis 3: Learners will elect to view immediate feedback more frequently than they elect to view delayed feedback

Study one found that 93.84 per cent of subjects had elected to view feedback after immediately responding to a question and 6.16 per cent of subjects had elected to view feedback at the end of the exercise. As discussed in the report of study three, an error in the code of the online delivery platform resulted in no data relating to feedback timing being captured for subjects studying in 2009. The same error resulted in no data being available during the initial analysis of study four. However, a retrospective recovery of data resulted in data being gathered for the following cohorts studying in the academic year 2010 to 2011.

Table 59: Cohorts for whom timing data were gathered in the academic year 2010 to 2011.

Test	Programme	Number of subjects
Accounting Primer Course	Executive Education	4
	Full time MBA	19
	IBSc Programme	34
	Joint Honours	20
	MSc Actuarial Finance	10
	MSc Finance	114
	MSc Int. Health Man.	32
	MSc Management	133
	MSc Risk Management	58
Pre-study Mathematics Course	Full time MBA	16
	IBSc Programme	12
	Joint Honours	9
Total		461

As shown in table 60 below, 21.91 per cent of subjects across these cohorts elected to view delayed feedback at the start of the exercise. However, the majority of these subjects changed this choice over the course of the activity. At the close of the activity, just 4.34 per cent of subjects had elected to view delayed feedback. This strongly supports the finding of study one and the hypothesis that subjects will elect to view immediate rather than delayed feedback.

Table 60: Feedback timing

	Delayed or Immediate	Frequency	%
Feedback position at start	D	101	21.91%
	I	360	78.09%
Changed mind*	D	345	74.84%
	I	116	25.16%
Feedback position at end	D	20	4.34%
	I	441	95.66%
Total		461	

4.5.4.6 Hypothesis 4: Learners' choice to view a feedback type is dependent on:

- a. Factor '#I5 Performance' (Whether a question is answered correctly or incorrectly)
- b. Factor '#E5 Question focus' (Whether a question tests recall or application of the study material)
- c. Factor '#I4 Level of expertise' (The level of the learner)
- d. Factor '#I6 Gender' (Whether the subject is male or female.)

As per studies one and three, a multilevel logistic regression was used to determine whether this hypothesis was supported. For this study, six factors were included in the analysis. The four factors stated in the hypothesis together with one variable indicating the item of TBI being taken (Accounting Primer course or Pre-study Mathematics course) and one variable indicating the programme on which a subject was studying (Full time MBA, IBSc Programme or Joint Honours). As described previously (See page 186), the individual question being studied was not included as a factor as this was considered to be captured by the variables 'Factor #I5 performance' and 'Factor #E5 question focus'.

First, variance inflation factors (VIFs) were calculated to identify the presence of collinearity between the independent predictor variables. The results showed low levels of collinearity and this as shown in table 61.

Table 61: Variance inflation factors (VIFs)

Variable	VIF	1/VIF
TBI	2.6	0.3852
Question focus	2.57	0.3889
IBSc Programme	1.22	0.8166
Joint Honours programme	1.2	0.8343
Gender	1.06	0.9477
Performance	1.05	0.9560
Learner level	1.04	0.9630
Mean VIF	1.53	

Next univariate and multivariate multilevel logistic regression analyses were performed for each feedback type. The results are shown in table 62 on the following page.

Table 62: Results of univariate and multivariate multilevel logistic regression analyses

Feedback type	Variable	Category	Univariable		Multivariable	
			Odds Ratio (95% CI)	P-value	Odds Ratio (95% CI)	P-value
NF	Performance	Incorrect	1		1	
		Correct	10.41 (4.98, 21.77)	<0.001	8.54 (4.01, 18.18)	<0.001
	Gender	F	1		1	
		M	2.84 (0.82, 9.88)	0.10	3.07 (0.81, 11.66)	0.099
	Question focus	Recall	1		1	
		Application	3.25 (2.30, 4.60)	<0.001	(0.29, 1.18)	0.133
	TBI	Accounting	1		1	
		Maths	5.34 (3.83, 7.45)	<0.001	7.48 (3.88, 14.43)	<0.001
	Programme	6	1		1	
		7	1.65 (0.40, 6.84)	0.49	1.85 (0.42, 8.15)	0.415
		8	1.41 (0.30, 6.65)	0.67	1.81 (0.35, 9.48)	0.482
	Learner level		0.99 (0.95, 1.04)	0.87	0.99 (0.94, 1.04)	0.633
KR	Performance	Incorrect	1		1	
		Correct	0.17 (0.10, 0.30)	<0.001	0.21 (0.12, 0.37)	<0.001
	Gender	F	1		1	
		M	0.45 (0.13, 1.57)	0.210	0.42 (0.11, 1.55)	0.192
	Question focus	Recall	1		1	
		Application	0.36 (0.26, 0.49)	<0.001	1.32 (0.72, 2.40)	0.372
	TBI	Accounting	1		1	
		Maths	0.24 (0.18, 0.33)	<0.001	0.21 (0.12, 0.37)	<0.001
	Programme	6	1		1	
		7	0.48 (0.12, 1.99)	0.315	0.43 (0.10, 1.89)	0.267
		8	0.67 (0.14, 3.16)	0.613	0.55 (0.11, 2.88)	0.482
	Learner level		1.01 (0.96, 1.05)	0.802	1.01 (0.96, 1.06)	0.609
KCR	Performance	Incorrect	1		1	
		Correct	0.01 (0.00, 0.01)	<0.001	0.01 (0.00, 0.01)	<0.001
	Gender	F	1		1	
		M	0.95 (0.71, 1.25)	0.700	1.46 (0.90, 2.38)	0.130
	Question focus	Recall	1		1	
		Application	1.19 (1.01, 1.41)	0.043	1.59 (1.10, 2.30)	0.014
	TBI	Accounting	1		1	
		Maths	0.65 (0.56, 0.76)	<0.001	0.69 (0.50, 0.97)	0.032
	Programme	6	1		1	
		7	1.17 (0.85, 1.62)	0.334	1.53 (0.88, 2.67)	0.129
		8	1.20 (0.84, 1.71)	0.312	1.38 (0.75, 2.55)	0.303
	Learner level		0.99 (0.98, 1.00)	0.0555	1.01 (0.99, 1.02)	0.591
IB	Performance	Incorrect	1		1	

	Gender	Correct	0.11 (0.09, 0.14)	<0.001	0.11 (0.09, 0.14)	<0.001
		F	1		1	
		M	1.16 (0.73, 1.86)	0.526	1.27 (0.74, 2.17)	0.380
	Question focus	Recall	1		1	
		Application	1.43 (1.12, 1.82)	0.004	1.41 (0.99, 2.01)	0.054
	TBI	Accounting	1		1	
		Maths	0.95 (0.77, 1.17)	0.623	0.98 (0.72, 1.34)	0.908
	Programme	6	1		1	
		7	0.96 (0.56, 1.65)	0.887	0.95 (0.52, 1.73)	0.869
		8	0.68 (0.37, 1.25)	0.216	0.63 (0.32, 1.25)	0.188
Learner level		1.00 (0.98, 1.01)	0.671	1.00 (0.98, 1.02)	0.848	
EX	Performance	Incorrect	1		1	
	Gender	Correct	0.07 (0.60, 0.92)	<0.001	0.07 (0.06, 0.09)	<0.001
		F	1		1	
		M	1.30 (0.87, 1.94)	0.203	1.44 (0.90, 2.33)	0.132
	Question focus	Recall	1		1	
		Application	2.01 (1.61, 2.51)	<0.001	1.80 (1.29, 2.51)	<0.001
	TBI	Accounting	1		1	
		Maths	1.23 (1.02, 1.47)	0.029	1.23 (0.92, 1.63)	0.163
	Programme	6	1		1	
		7	1.19 (0.76, 1.88)	0.452	1.29 (0.76, 2.19)	0.342
		8	0.69 (0.41, 1.15)	0.158	0.65 (0.36, 1.20)	0.169
	Learner level		1.01 (0.99, 1.020)	0.417	1.02 (1.00, 1.04)	0.095
	EI1	Performance	Incorrect	1		1
Gender		Correct	0.32 (0.22, 0.46)	<0.001	0.33 (0.23, 0.48)	<0.001
		F	1		1	
		M	1.94 (0.94, 4.00)	0.073	1.91 (0.89, 4.09)	0.094
Question focus		Recall	1		1	
		Application	0.84 (0.60, 1.18)	0.311	1.04 (0.64, 1.69)	0.881
TBI		Accounting	1		1	
		Maths	0.68 (0.50, 0.93)	0.017	0.73 (0.46, 1.15)	0.176
Programme		6	1		1	
		7	0.57 (0.24,1.29)	0.176	0.60 (0.26, 1.41)	0.243
		8	0.46 (0.18, 1.15)	0.097	0.48 (0.19, 1.26)	0.137
Learner level			1.00 (0.97, 1.02)	0.824	0.99 (0.97, 1.02)	0.706
EI2		Performance	Incorrect	1		1
	Gender	Correct	0.60 (0.40, 0.89)	0.011	0.63 (0.42, 0.94)	0.024
		F	1		1	
		M	1.46 (0.75, 2.82)	0.261	1.38 (0.69, 2.76)	0.370
	Question focus	Recall	1		1	
		Application	0.75 (0.53, 1.04)	0.086	1.05 (0.65, 1.71)	0.843
	TBI	Accounting	1		1	

	Maths	0.64 (0.47, 0.88)	0.005	0.64 (0.41, 1.02)	0.059
Programme	6	1		1	
	7	0.75 (0.35, 1.62)	0.469	0.81 (0.37, 1.77)	0.604
	8	0.83 (0.36, 1.91)	0.660	0.92 (0.39, 2.17)	0.845
Learner level		1.02 (0.99, 1.05)	0.212	1.02 (0.99, 1.05)	0.260

The odds ratio for factors significant to a P level of less than 0.05 are summarized in the following table.

Table 63: Summary of results for factors significant to a p level of less than 0.05

	Performance		TBI		Ques. focus	
	OR	P	OR	P	OR	P
NF	8.54	<0.001	7.48	<0.001		
KR	0.21	<0.001	0.21	<0.001		
KCR	0.01	<0.001	0.69	0.032	1.59	0.014
IB	0.11	<0.001				
EX	0.07	<0.001			1.8	<0.001
EI1	0.33	<0.001				
EI2	0.63	<0.001				

As in study three, factor ‘#I5 Performance’ was found to be a significant factor for all feedback types. Having answered a question correctly subjects were significantly less likely to view feedback types KR (OR=0.21), KCR (OR=0.01), IB (OR=0.11), EX (OR=0.07), EI1 (OR=0.33) and EI2 (OR=0.63) and therefore more likely not to view feedback (OR=8.5)

In study three, the item of TBI, was found to be the most significant factor. In this study the effect was not as pronounced. Subjects studying the Pre-Study Mathematics course were less likely to view KR type feedback (OR=0.21) and KCR type feedback and therefore more likely not to view feedback (OR=7.48). There was no effect of TBI for the remaining feedback types.

Study three found that factor ‘#E5 Question focus’, whether the question being answered tests application or recall of the material, studied had a significant effect on whether subjects elected to view feedback types KCR, IB, EX, EI1 and EI2. Subjects were more likely to view these feedback types if a question tested application of the material studied. In this study similar positive effects were found for feedback types KCR (OR=1.59) and EX (OR=1.80) however there were no significant effects for feedback types IB, EI1 or EI2.

No significant effects were found for factors ‘#I4 Level of expertise’, ‘#I6 Gender’ or the degree programme on which a subject was studying.

4.5.4.7 Hypothesis 5: *The majority of learners will revisit a previously viewed question*

As described in studies one and three, the online system tracks subjects' pathways through the questions and this reveals whether subjects chose to revisit a question. In study one, 86.30 per cent of students revisited at least one question over the course of the activity and on average, subjects revisited 4.17 questions. In addition, subjects elected to view an additional feedback item on 42.19 per cent of revisits. In study three, 42.85 per cent of a sample of subjects revisited at least one question and these subjects revisited 1.42 questions on average when studying sections of the Accounting Primer course (The balance sheet) and Pre-Study Mathematics course (Basic Maths 2) each of which contained six questions. 19.84 per cent of students both revisited at least one question and viewed an additional feedback item when they did so.

In study four, the pathways of subjects from the following cohorts of students were analysed to see if this results was supported.

Table 64: Cohorts selected for analysis for hypothesis 5

Cohort	Year	Number of students	Test	Section
Full time MBA	2010	38	Accounting Primer	The Balance sheet (Six questions)
IBSc Programme	2010	39	Pre-study Mathematics	Basic Maths 2 (Six questions)

As shown in table 65 below, 37.69 per cent of subjects revisited at least one question. Across the whole cohort subjects revisited 1.38 questions on average. 20.78 per cent of students both revisited at least one question and viewed an additional feedback item when they did so.

Table 65: Subjects revisiting questions

	Accounting Primer course section	Pre-study Mathematics course section	Mean across the two sections
Proportion of students who revisited at least one question	0.3947	0.3590	0.3769
Proportion of students who revisited question and viewed additional feedback	0.2105	0.21	0.2078
Mean number of questions revisited across the cohort	1.61	1.15	1.38
Mean number of additional feedback items viewed across the cohort	0.55	0.46	0.51

Hypothesis 5 is not supported. The majority of subjects did not revisit a question.

4.5.4.8 *Hypothesis 6: Learners' use of feedback will not change significantly over the course of the activity*

As described in study three (See page 200), the Pre-study Mathematics tests and the Accounting Primer tests are divided into smaller tests comprising between five and six questions. There are insufficient questions in these tests to reveal any patterns over the course of the test. Hypothesis 6 is neither supported nor not supported.

4.5.5 Analysis and conclusions

4.5.5.1 *Subjects viewed feedback extensively.*

Studies one and three found that subjects chose to view feedback more often than they did not. Study one found that subjects elected to view at least one feedback type for 85.19 per cent of the questions they answered. In study three, this figure rose to 92.50 per cent and in study four this rose again to 94.67 per cent. It is apparent from these studies that, if provided, students will view feedback extensively.

4.5.5.2 *The existence of a threshold does not hold for all items of TBI*

In studies one, three and four subjects elected to view feedback types KR and KCR more frequently than types IB, EX, EI1 and EI2. In study one, this difference was unambiguous and this provided support for the existence of a threshold, at the level of KCR type elaborative feedback, beyond which the study of feedback becomes either ineffective or inefficient (Phye 1979; Kulhavy, White et al. 1985; Dempsey and Litchfield 1993; Pridemore and Klein 1995). In study three the difference in mean α values was not as striking and did not give the appearance of a threshold. This was also the case in study four.

In particular, as in study three, the value for EX for the Pre-study Mathematics course (0.1536) was noticeably closer to the value for KCR (0.1794). In other words subjects elected to view an explanation of how to answer a question correctly almost as often as they elected to view the correct answer. The results of the investigation into hypothesis four show the item of TBI being studied had no effect on whether subjects elected to view the elaborative feedback types beyond KCR. However, a strong effect was found for factor '#E5 Question focus' on feedback type EX. Subjects were more likely to click on this feedback type when a question tested application rather than recall of the material. All questions in the Pre-study Mathematics course tested application of the material and this provides an explanation for the high mean α values for feedback type EX.

For the Pre-study Mathematics course, subjects' actions suggest they disagree with the hypothesis that feedback beyond the level of KCR is always either ineffective or inefficient.

4.5.5.3 *Subjects held a preference for feedback that focuses on the knowledge gap*

A hypothesis emerged in study one that, following an incorrect response, subjects hold a preference for elaborative feedback beyond KCR that focuses on any identified 'knowledge gap'.

Study two provided some support for this hypothesis. In the analysis that followed, this hypothesis was combined with the conclusions of study two and the constructivist notion of a ‘perturbation’ to form the basis of a model of the process through which learners engage with TBI feedback. In this model, learners first elect to view KR type feedback in order to receive affirmation that their response was correct. Following this, learners proceed to view the elaborative feedback types if a knowledge gap has been identified. Learners seek the most efficient means of resolving this gap, that is they seek to view the minimum load of TBI feedback required. Subjects first view KCR type feedback as this is often sufficient to close the knowledge gap and is considered most efficient. If required, subjects then view the remaining elaborative feedback types in the order of their specificity on the knowledge gap, namely EX type feedback then IB type feedback and then the extra-instructional feedback types EI1 and EI2.

If subjects do view the remaining elaborative feedback types in the order of their specificity on the knowledge gap then the frequency at which learners elect to view these feedback types should reflect this order. Study three provided some support for this hypothesis. When mean α values for the three items of TBI were considered together, this order was that hypothesized. However, an examination of mean α values by item of TBI revealed an exception, the mean α value for IB type feedback was higher than that for EX type feedback for questions within the Accounting Primer course. However, an explanation for this discrepancy was detailed in the ‘Analysis and conclusions’ section in the report of study three (See page 202) and the conclusion reached was that the results of study three did provide support for the hypothesis the subjects hold a preference for feedback that focuses on the knowledge gap.

Should this hypothesis hold for study four then we should expect slightly different results for the Pre-Study mathematics courses and the Accounting Primer course. For the former the mean α value for KCR type feedback should be the highest followed by EX then IB and then the extra-instructional feedback types EI1 and EI2. For the Accounting Primer course we should find that the mean α value for KCR type feedback should be the highest followed by EX then IB and then types EI1 and EI2.

As shown in table 66 below, the results match those hypothesised with one exception. For the Accounting Primer course, the mean α value for IB type feedback is lower than that for EX type feedback.

Table 66: Mean α values by item of TBI

	No feedback	KR	KCR	IB	EX	EI1	EI2
All	0.0533	0.9397	0.2082	0.1099	0.141	0.0496	0.048
Accounting Primer Course only	0.0235	0.9653	0.242	0.1106	0.1261	0.0489	0.0521
Pre-study Mathematics only	0.0787	0.9178	0.1794	0.1092	0.1536	0.0503	0.0445

To investigate this issue further, data were gathered for all cohorts who had studied the Accounting Primer course over the course of this research. The total number of subjects is 1,123 across 21 cohorts of students. As shown in

table 67 below, the mean α value for IB type feedback is higher than that for EX type feedback for 19 of the 21 cohorts. For one cohort of students the values are equal and for one cohort of students the mean α value for EX type feedback is higher than that for IB type feedback, the 'IBSc Programme' 2010.

Table 67: Mean α values –Cohorts studying Accounting Primer course in 2009 & 2010

Cohort		n	No feedback	KR	KCR	IB	EX	EI1	EI2
2009	Best	19	0.04	0.95	0.44	0.35	0.22	0.21	0.27
	Full time MBA	58	0	1	0.17	0.08	0.07	0.03	0.04
	IBSc Programme	51	0.04	0.96	0.27	0.11	0.11	0.06	0.05
	Joint Honours	23	0.03	0.97	0.29	0.18	0.08	0.08	0.1
	MSc Actuarial Finance	12	0.03	0.88	0.35	0.21	0.09	0.06	0.03
	MSc Finance	101	0.05	0.93	0.27	0.16	0.13	0.1	0.1
	MSc Int Health Man	43	0.04	0.96	0.25	0.16	0.11	0.12	0.13
	MSc Management	118	0.02	0.97	0.24	0.12	0.08	0.06	0.05
	MSc Risk Man	66	0.04	0.96	0.27	0.21	0.14	0.14	0.12
	MSc Civil Env Eng	13	0.03	0.97	0.39	0.32	0.16	0.26	0.28
2010	Executive Education	24	0.01	0.99	0.26	0.12	0.11	0.11	0.08
	Executive MBA	22	0.05	0.95	0.17	0.13	0.12	0.05	0.03
	Full time MBA	69	0.05	0.95	0.23	0.14	0.12	0.08	0.08
	IBSc Programme	67	0.01	0.97	0.23	0.08	0.11	0.03	0.03
	Joint Honours	38	0	0.99	0.23	0.08	0.07	0.04	0.06
	MSc Actuarial Finance	14	0.1	0.9	0.32	0.16	0.13	0.05	0.06
	MSc Finance	135	0.05	0.95	0.28	0.18	0.13	0.09	0.09
	MSc Int Health Man	39	0.05	0.94	0.34	0.22	0.16	0.12	0.11
	MSc Management	145	0.04	0.95	0.28	0.17	0.14	0.09	0.06
	MSc Risk Man	77	0.02	0.96	0.26	0.14	0.12	0.09	0.07
	Weekend EMBA	39	0.02	0.98	0.26	0.16	0.12	0.1	0.07

Thus, the anomaly in study four arises arising from data relating to one exceptional group, the 'IBSc Programme' in 2010. Across all subjects the mean α value for IB type feedback was 0.17. For EX type feedback the mean α value was 0.12.

Given this discovery, there was also a risk that the result for the pre-Study Mathematics course had been subject to a similar outlier. However, as shown in table 68 below, the hypothesized order held for nine of the ten cohorts for whom data were collected. The single cohort for whom the hypothesis did not hold consisted of one student and can therefore be considered an outlier.

Table 68: Mean α values –Cohorts studying Pre-study Mathematics course in 2009 & 2010

Cohort		n	No feedback	KR	KCR	IB	EX	EI1	EI2
2009	Best								
	Full time MBA	221	0.08	0.89	0.17	0.1	0.14	0.03	0.02
	Joint Honours	298	0.09	0.9	0.21	0.12	0.15	0.04	0.04
	MSc Management	686	0.08	0.89	0.14	0.12	0.14	0.07	0.04
2010	DL MBA	1	0	1	0.27	0	0.55	0.18	0.18
	Executive MBA	19	0.01	0.99	0.21	0.16	0.22	0.06	0.05
	Full time MBA	80	0.07	0.92	0.17	0.11	0.14	0.07	0.06
	IBSc Programme	69	0.08	0.91	0.19	0.12	0.18	0.05	0.05
	Joint Honours	47	0.04	0.96	0.17	0.09	0.1	0.02	0.04
	MSc Management	29	0.19	0.77	0.21	0.03	0.18	0.02	0.01
	Weekend EMBA	30	0.07	0.93	0.17	0.12	0.17	0.02	0.02

The conclusion reached is that subjects did hold a preference for feedback that focuses on the knowledge gap in study four.

4.5.5.4 *Patterns in subjects' behaviour*

When investigating the distribution of α values across the subject cohort, studies one and three found a common pattern. A large majority of subjects elected to view feedback type KR for virtually all of the questions. For the elaborative feedback types, a significant proportion of the subjects did not elect to view these feedback types, the majority viewed them rarely but a minority of students elected to view these feedback types across a range of frequencies. Thus, when plotted on a graph, the distributions of α values across the cohort contained a marked positive skew for the KCR, IB, EX, EI1 and EI2 elaborative feedback types.

In study one, differences in subjects' performance largely accounted for the different frequencies at which subjects elected to view KCR type feedback and partly explained the same result for feedback types IB and EX. In study three, subjects' performance largely accounted for the differences in subject behaviour for feedback types KCR, IB, EX, EI1 and EI2.

The results of this study closely resemble those of study three. Again, a large majority of subjects elected to view feedback type KR for virtually all of the questions. And a significant positive skew appeared in the distribution of α values for the elaborative feedback types. As in study three, the variation across these distributions was largely, but not totally, accounted for by differences in subjects' performance.

4.5.5.5 Factors affecting subjects' choice to view a feedback type

The effect of Factor #I5 Performance

Studies one and three found that performance had a significant and sizable effect on whether a subject elected to view the feedback types. As in study three, this study found that performance was found to be a significant factor for all six feedback types. Having answered a question correctly subjects were significantly less likely to view feedback types KR (OR=0.21), KCR (OR=0.01), IB (OR=0.11), EX (OR=0.07), EI1 (OR=0.33) and EI2 (OR=0.63) and therefore more likely not to view feedback (OR=8.54).

As mentioned in the previous section 'Patterns in subjects' behaviour', the differing levels in performance in the tests across the subject cohort also explained a significant proportion of the variation in the distribution of α values across the cohort.

The effect of Factor #E5 Question focus

In study three, factor '#E5 Question focus' was found to be significant factor in determining the frequency at which subjects elected to view feedback types KCR, IB, EX, EI1 and EI2 and also whether subjects chose not to view feedback. Subjects were significantly more likely to view these feedback types if a question tested application of the material studied.

Study four found similar positive effects for feedback types KCR and EX however there were no significant effects for feedback types IB and EI1. A univariate analysis did find an effect for IB type feedback type (OR=1.43, $p=.004$) however in the multivariate analysis this effect proved not to be significant (OR=1.41, $p=.054$). However, it may be the case that a larger sample size, as in study three, would find this effect to be significant.

The odds ratios for the significant effects of factor '#E5 question focus' in studies three and four are shown in the following table. In each case, the effect was $p < 0.001$.

Table 69: Odds ratios for the significant effects of factor '#E5 question focus'

	No feedback	KR	KCR	IB	EX	EI1	EI2
Study three	0.78 (.0038)		1.57 ($<.001$)	1.64 ($<.001$)	2.28 ($<.001$)	1.69 ($<.001$)	1.48 ($<.001$)
Study four			1.59 (.014)		1.80 ($<.001$)		

From studies three and four we can conclude that learners are more likely to view KCR and EX type feedback if questions test application as opposed to recall of the course material. Question focus may also affect the frequency in which learners elect to view IB, EI1 and EI2 type feedback however this needs to be subject to further investigation.

The effect of Factor #I4 Level of expertise

Three studies identified in the literature review indicate that learners' ability in a subject may impact on the effectiveness of feedback (Roper 1977; Butler and Winne 1995; Handley 2002) and this formed the basis of the hypothesis that subjects' level of expertise affects learners' choice to view the feedback types.

In this study, level of expertise was termed factor 'I4 Level of expertise' and measured using subjects' score in the final assessment of the items of TBI. However, no significant effects for this variable factor were found.

The effect of Factor #I6 Gender and the degree programme on which a subject was studying

No significant effects were found for either factor 'I6 Gender' or the degree programme on which a subject was studying.

The effect of the TBI material

In study three 'TBI', in other words the item of TBI being studied, was found to be the most significant of the factors included in the test of hypothesis four. In study four the effect was noticeably smaller. Subjects studying the Pre-Study Mathematics course were less likely to view KR type feedback (OR=0.21) and KCR type feedback and therefore more likely not to view feedback (OR=7.48) however there was no effect of 'TBI' for the remaining feedback types.

Thus the large effect for 'TBI' observed in study three seems due to the presence of the Ethics and Strategy exercise. One explanation is that the large effect is due to the content of this exercise. It contains different questions, feedback and subject matter to the other items of TBI. However, an alternative explanation is that the conditions in which the subjects took this activity were different. Subjects studying the Accounting Primer and Pre-study Mathematics courses did so at a time and place of their choosing. However subjects took the Ethics & Strategy exercise together in a computer laboratory and with a time limit. This alternative explanation is given support by the interviews in study two during which the subjects interviewed stressed the effect of the time limit in particular on their decision to view the feedback types.

4.5.5.6 Subjects prefer immediate feedback

The literature review revealed an on-going and inconclusive debate as to whether feedback should be provided immediately after a response or following a delay. In study one, subjects expressed a clear preference for immediate feedback. 137 of the 146 subjects elected to view feedback immediately following their response to a question. This result was repeated in study four. At the end of the activity 441 out of 461 subjects had elected to view immediate feedback.

Subjects in both these studies expressed a clear preference for immediate rather than delayed feedback.

4.5.5.7 *Subjects revisit questions and feedback*

In study one, 86.30 per cent of students revisited at least one question over the course of the activity and on average, subjects revisited 4.17 questions. In addition, subjects elected to view an additional feedback item on 42.19 per cent of revisits. This led to the hypothesis that the majority of students will revisit a previously viewed question.

This hypothesis was not supported in study three and neither was it supported in study four. 37.69 per cent of subjects revisited at least one question, subjects revisited 1.38 questions on average and 20.78 per cent of students both revisited at least one question and viewed an additional feedback item when they did so.

This result is similar to that of study three and again the marked difference with the results of study one is explained by the difference in the length of the tests. The test in the Ethics and Strategy exercise from study one consisted of 50 questions while the tests in the sections of the Accounting Primer course and Pre-study Mathematics course consisted of six questions.

However, as per study three, the result that 37.69 per cent of subjects did choose to revisit a question is of interest.

4.5.6 **Summary**

The aim of this study was to test the results of studies one and three with additional cohorts of students and also to test the additional hypothesis that factor '#I4 level of expertise' affects learners' choice to view a feedback type.

As in studies one and three, hypothesis one was supported. Subjects viewed feedback extensively, choosing to view at least one feedback type for 94.67 per cent of the questions they answered.

As in study three, the results of the test of hypothesis two suggest that the existence of a feedback threshold at the level of KCR type feedback may depend on the material being studied. For the Pre-study Mathematics course, subjects' actions suggest they disagree that viewing the elaborative feedback types beyond KCR type feedback was always either ineffective or inefficient.

Further evidence emerged from this study to support the hypothesis that learners hold a preference for feedback that focuses specifically on the knowledge gap with the pattern of results closely resembling those in study three.

As in study one, the result of the test of hypothesis three was unequivocal, subjects demonstrated a clear preference for immediate feedback.

The test of hypothesis four largely supports the results of study three. Again, performance was found to be a significant factor for all six feedback types. Having answered a question correctly subjects were significantly less likely to view feedback types KR, KCR, IB, EX, EI1 and EI2. However, results for factor '#E5 Question focus' differ. Study three found that this factor to have a significant effect on whether subjects elected to view feedback types KCR, IB, EX, EI1 and EI2. Subjects were more likely to view these feedback types if a question tested

application of the material studied. In this study, similar significant positive effects were found for feedback types KCR and EX however there were no significant effects for feedback types IB, EI1 or EI2. In addition, the test of hypothesis four found no significant effects for factors ‘#I4 Level of expertise’, ‘#I6 Gender’ or the degree programme on which a subject was studying.

The patterns of behaviour observed this study closely resemble those of study three and again support the findings of study one. Virtually all subjects elected to view feedback type KR for virtually all of the questions. For the elaborative feedback types, the distribution of α values across the cohort contained a marked positive skew. For KCR type feedback this variation was predominantly explained by differences in performance across the cohort. However, performance was not able to totally explain the variety in subjects’ behaviour for feedback types IB, EX, EI1 and EI2.

As per study three, fewer students chose to revisit questions than in study one. 37.69 per cent of subjects did so as opposed to 86.30 per cent in study one. As before, an explanation is that this is due to the tests comprising fewer questions.

5. SUMMARY OF RESULTS

This chapter summarizes the results of studies one to four starting with the results of hypothesis 1.

5.1.1 Hypothesis 1: Learners will elect to view feedback more frequently than they do not

Support for this hypothesis is provided in studies one, three and four. Table 70 below displays the mean and median α values for each feedback type together with the standard deviation for all subjects contributing data to this series of studies. To illustrate, this table shows that on average subjects clicked on feedback type KR for 91.11 per cent of the questions they answered.

Table 70 shows that subjects did not view feedback for 7.28 per cent of the questions they answered and therefore viewed feedback for 92.72 per cent of the questions they answered. Hypothesis 1 is supported.

Table 70: α values for each feedback type for all subjects across this series of studies

	No feedback	KR	KCR	IB	EX	EI1	EI2
N	2066	2066	2066	2066	2066	2066	2066
Mean	0.0728	0.9111	0.2041	0.1241	0.1314	0.0597	0.0504
Median	0.0000	1.0000	0.1670	0.0000	0.0000	0.0000	0.0000
Std. Deviation	0.2221	0.2465	0.2434	0.2225	0.2323	0.1670	0.1500

The following tables show the same statistics by study and item of TBI. Hypothesis 1 is supported in each case.

Table 71: α values for all subjects across studies one to four by cohort and item of TBI

Ethics and Strategy exercise - 2008 cohort

	No feedback	KR	KC	IB	EX	EI1	EI2
N	146	146	146	146	146	146	146
Mean	0.1481	0.8447	0.3223	0.0655	0.0889	0.0203	0.0191
Median	0.0000	1.0000	0.3200	0.0200	0.0200	0.0000	0.0000
Std. Deviation	0.3259	0.3336	0.2124	0.1194	0.1743	0.0573	0.0642

Pre-Study Mathematics course - 2009 cohort

	No feedback	KR	KCR	IB	EX	EI1	EI2
N	1205	1205	1205	1205	1205	1205	1205
Mean	0.0843	0.8942	0.1638	0.1199	0.1439	0.0529	0.0373
Median	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Std. Deviation	0.2411	0.2703	0.2527	0.2355	0.2581	0.1684	0.1347

Accounting Primer course - 2009 cohort

	No feedback	KR	KCR	IB	EX	EI1	EI2
N	504	504	504	504	504	504	504
Mean	0.0316	0.9588	0.2642	0.1570	0.1100	0.0917	0.0918
Median	0.0000	1.0000	0.2000	0.0910	0.0000	0.0000	0.0000
Std. Deviation	0.1310	0.1538	0.2276	0.2271	0.1931	0.1874	0.1948

Accounting Primer Course - 2010 cohort

	No feedback	KR	KCR	IB	EX	EI1	EI2
N	97	97	97	97	97	97	97
Mean	0.0235	0.9653	0.2420	0.1106	0.1261	0.0489	0.0521
Median	0.0000	1.0000	0.1820	0.0000	0.0590	0.0000	0.0000
Std. Deviation	0.1009	0.1415	0.2056	0.1930	0.2158	0.1473	0.1507

Pre-study Mathematics - 2010 cohort

	No feedback	KR	KCR	IB	EX	EI1	EI2
N	114	114	114	114	114	114	114
Mean	0.0787	0.9178	0.1794	0.1092	0.1536	0.0503	0.0445
Median	0.0000	1.0000	0.1600	0.0720	0.1200	0.0000	0.0000
Std. Deviation	0.2053	0.2137	0.1568	0.1562	0.1576	0.1457	0.1177

5.1.2 Hypothesis 2: Learners will elect to view KCR and KR type feedback more frequently than they elect to view either EX, IB or EI type feedback.

Support for this hypothesis is provided in studies one, three and four. Subjects generally elected to view KR type feedback, they did so on 91.11% of the questions they answered. In addition, subjects generally elected to view KCR type feedback if they had responded incorrectly. Table 72 shows how frequently subjects elected to view KCR type feedback following an incorrect response.

Table 72: Use of KCR type feedback - Incorrect responses only

	Frequency	%
Did not view KCR type feedback	1,908	16.81%
Did view KCR type feedback	9,441	83.19%
Total	11,349	100.00%

In addition, subjects elected to view both of these feedback types more frequently than they elected to view the EX, IB or EI feedback types. When all subjects in this series of studies are considered, the mean α value for IB type feedback is 0.1241 meaning that, on average, subjects elected to view the relevant section of the original instruction for 12.41 per cent of the questions they answered. The mean α value for EX type feedback is 0.1100 and for EI1 and EI2 type feedback is 0.0917 and 0.0918 respectively. The α value for all four feedback types is lower than that of both KR (0.9111) and KCR (0.2041) type feedback and hypothesis 2 is therefore supported.

This hypothesis is also supported when looking at the data by study and by item of TBI. In each case, subjects elected to view KCR and KR type feedback more frequently than they elect to view either EX, IB or EI type feedback.

It was of note that the α value of EX type feedback for the Pre-study Mathematics tests in studies three and four was noticeably closer to the value for KCR. This is discussed further in the 'Discussion' chapter that follows (See page 251).

5.1.3 Hypothesis 2.1: For elaborative feedback types beyond KCR: Learners will elect to view EX most frequently followed by IB and then EI type feedback.

Support for this hypothesis varied across the items of TBI. The hypothesis was fully supported in study one which found that subjects viewed EX type feedback most frequently followed by IB, EI1 and then EI2 type feedback. This hypothesis was supported in study three when subjects studied the Ethics and Strategy exercise and the Pre-study Mathematics course but not when studying the Accounting Primer course. This is discussed in the following

chapter ‘Discussion’ (See page 252). The hypothesis was supported in study four for subjects studying both the Pre-study Mathematics course and the Accounting Primer course.

Hypothesis 2.1 is supported when all cohorts across this research are considered. Table 73 below shows the mean α values for feedback types IB, EX, EI1 and EI2 for all cohorts in the studies comprising this research. Subjects elected to view EX (0.1314) most frequently, followed by IB (0.1241), EI1 (0.0597) and then EI2 (0.0504) type feedback.

Table 73: Mean α values for feedback types IB, EX, EI1 and EI2 for all cohorts

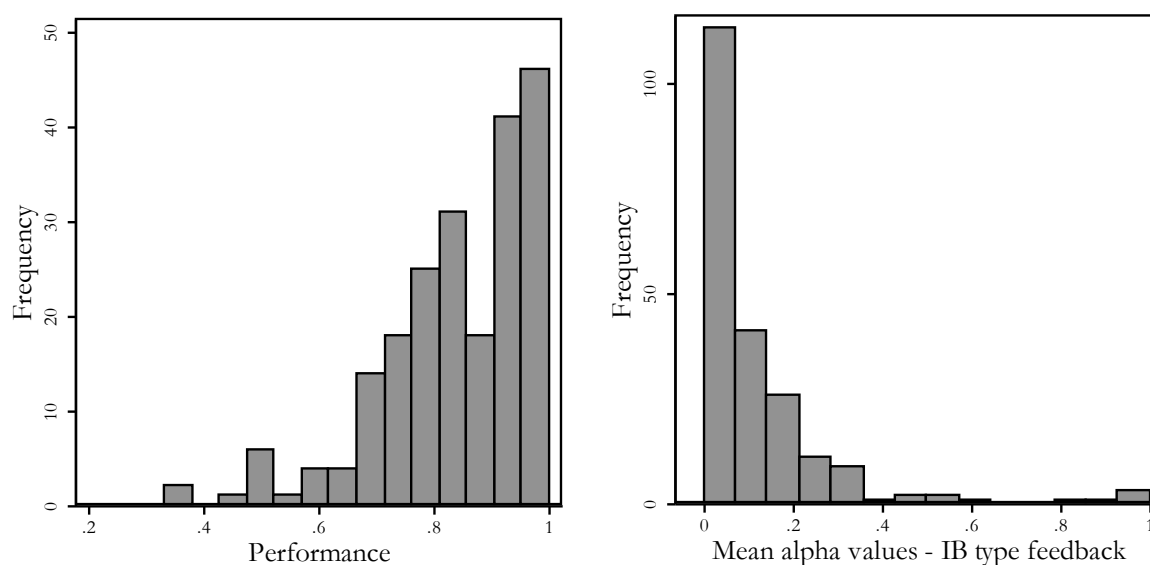
		IB	EX	EI1	EI2
All subjects		0.1241	0.1314	0.0597	0.0504
2008	Ethics and Strategy Exercise	0.0655	0.0889	0.0203	0.0191
2009	Accounting Primer	0.1570	0.1100	0.0917	0.0918
2009	Pre-study Mathematics	0.1199	0.1439	0.0529	0.0373
2010	Accounting Primer	0.1106	0.1261	0.0489	0.0521
2010	Pre-study Mathematics	0.1092	0.1536	0.0503	0.0445

5.1.4 Hypothesis 2.2: The distribution of α values for elaborative feedback types will contain a marked positive skew

Studies one, three and four found that the distribution of α values for the KCR, IB, EX, EI1 and EI2 elaborative feedback types contained a marked positive skew. A significant proportion of the subjects did not elect to view these feedback types, the majority viewed them rarely and a minority of students elected to view these feedback types across a range of frequencies.

Plotting the distributions of α values for these feedback types as histograms revealed that the subsequent distribution patterns reflected the distribution pattern for student performance, particularly for those subjects studying the Accounting Primer and Pre-study Mathematics courses. This is shown in figure 55 below which presents the distribution of performance scores and the distribution of mean α values for the subjects in study four.

Figure 55: Performance scores and mean α values for IB feedback type in study 4
(Accounting & Mathematics courses)



For both studies three and four this effect was confirmed in the test of hypothesis four which found performance to be a significant factor in determining how frequently subjects elected to view all feedback types.

For students studying the Ethics and Strategy exercise in study one this effect was not as marked. Performance was found to be a significant factor for feedback types KCR, IB and EX however the size of the effect for the latter two types was not sufficient to explain much of the pattern observed in the distribution of α values. This is discussed further in the following chapter 'Discussion' (See page 255).

5.1.5 Hypothesis 3: Learners will elect to view immediate feedback more frequently than they elect to view delayed feedback

Studies one and four provide strong support for this hypothesis. Table 74 below shows the combined data from studies one and four relating to subjects' preference for the timing of feedback.

Table 74: Preference for the timing of feedback

	Delayed or Immediate	Frequency	%
Feedback position at start	D	171	28.17%
	I	436	71.83%
Changed mind	D	398	65.57%
	I	209	34.43%
Feedback position at end	D	29	4.78%
	I	578	95.22%
Total		607	

Prior to starting the exercise, subjects were given the choice of viewing feedback after responding to a question (immediate feedback) or viewing feedback at the end of the exercise (delayed feedback). The table shows that at

this stage 28.17 per cent of the cohort elected to view delayed feedback and 71.83 per cent elected to view immediate feedback.

After starting the exercise subjects were able to change this setting. The table shows that 65.57 per cent of subjects who had initially selected delayed feedback did so and this had a significant effect. 95.22 per cent of subjects ended the exercise electing to view feedback after responding to a question and therefore just 4.78 per cent of subjects electing to view feedback at the end of the exercise.

Subjects expressed an unambiguous preference for immediate rather than delayed feedback and hypothesis 3 is supported.

5.1.6 Hypothesis 4: Factors that affect whether learners choose to view a feedback type

The literature review identified nine factors that potentially affect whether a learner elects to view a feedback type. Seven of these were classified as being internal to the learner and two were classified as being external, or environmental, variables. Study two identified an additional four external factors resulting in a total of 13 potential factors being identified. research.

Table 75: Factors that potentially affect whether learners choose to view a feedback type

Internal factors	External factors
Factor #I1 Learner expectations	Factor #E1 Assessment
Factor #I2 Response certitude	Factor #E2 Feedback level
Factor #I3 Interpretation of task	Factor #E3 Time limit
Factor #I4 Level of expertise	Factor #E4 Location
Factor #I5 Performance	Factor #E5 Question focus
Factor #I6 Gender	Factor #E6 Subject matter
Factor #I7 Error type	

Five of these factors were investigated in studies one, three and four using logistic regression analyses. These five factors are shown in the following table.

Table 76: The five factors investigated via logistic regression

Internal factors	External factors
Factor #I4 Level of expertise	Factor #E1 Assessment
Factor #I5 Performance	Factor #E5 Question focus
Factor #I6 Gender	

10 factors were investigated explicitly in the semi-structured interviews that comprised study two. These are shown in table 77.

Table 77: The ten factors investigated in the semi-structured interviews

Internal factors	External factors
Factor #I1 Learner expectations	Factor #E1 Assessment
Factor #I2 Response certitude	Factor #E3 Time limit
Factors #I3 Interpretation of task	Factor #E4 Location
Factor #I4 Level of expertise	Factor #E5 Question focus
Factor #I5 Performance	Factor #E6 Subject matter

The results for each factor are now summarized.

Factor #I1 Learner expectations

The literature review identified learners' expectations relating to feedback to be a factor potentially affecting whether they choose to view a feedback type and the ten interviews comprising study two provided support for this hypothesis. Handley (2003: 208) found that learners' expectations of feedback arise from four factors:

1. Learners' **'style of Engagement'** with the instructional activity
Handley defines seven distinct styles ranging from 'mindful learning' to 'going through the motions'.
2. Learners' **'sense of inquiry'** arising from the task
This refers to learners' level of interest in the activity.
3. Learners' **'assessment of personal risk'** arising from the task
This refers to learners' assessment of their ability to deal with the real life scenarios presented in Handley's study together with the implications of failure.
4. Learners' **generic expectations of feedback**
Broadly speaking, learners generally expect feedback to verify, correct, and explain.

In study two subjects stated both 'style of engagement' and 'sense of enquiry' to be factors determining whether they elected to view the feedback types. Subjects revealed they had adopted at least four of the styles of engagement from Handley's taxonomy when taking the activity. These being 'Playing the game', 'Going through the motions', 'Getting through it' and 'Mindful learning'. Subjects stated their style of engagement to be a determining factor when electing to view a feedback type.

'Sense of inquiry' relates to the degree of interest learners hold in an activity. In study two, subjects used a variety of terms to refer to their sense of inquiry and, again, stated this to affect their choice to view a feedback type. These

terms included 'like', 'curiosity' and required degree of 'satisfaction'. This sense of inquiry operated at three different levels, the course level, the activity level and the question level.

Study two thus provided support for the hypothesis that factor '#1 learner expectations' affects learners' decision to view a feedback type. This factor was not investigated in studies one, three and four. However these studies did observe variation in the α values across the subject cohort not completely explained by differences in subjects' performance. As discussed in the report of study two (See page 175), factors '#I1 learner expectations' and '#I3 interpretation of task' offer some explanation for this variation.

In addition to identifying differences in α values, studies one, three and four also found that subjects elected to view different combinations of feedback types across different questions. Subjects generally viewed KR type feedback for all questions and then KCR type feedback when they had responded incorrectly. Subjects did not view the remaining feedback types for the majority of questions but did for a minority. Study two found factor '#I1 learner expectations' to be one factor determining whether a subject will view these additional feedback types along with factors '#I2 response certitude' and '#I3 interpretation of task' (See page 173).

Factor #I2 Response certitude

The ten interviews comprising study two strongly support the hypothesis that factor '#I2 Response certitude' affects learners decision to view a feedback type.

The term 'response certitude' was proposed by Kulhavy and Stock (1989) who found that the degree to which learners are certain that a response is correct affects how long students spend on instructional feedback following an incorrect response. For example, if learners are certain they have answered a question correctly but are informed they responded incorrectly then they will spend more time on feedback than if they had been less certain.

In study two, seven of the ten subjects stated response certitude to be a significant factor when choosing to move beyond KR and KCR type feedback to view the IB, EX, EI1 or EI2 elaborative feedback types. All seven subjects stated they elected to view this elaborative feedback when they held a high level of certainty that their answer had been correct.

To a lesser degree, subjects stated they would elect to view the elaborative feedback types when they had a low certainty that their response was correct, when they had guessed. Study one found that, while subjects were more likely to view feedback types KCR, IB and EX following an incorrect response, they also elected to view these feedback types following a correct response. The presence of guessing offers one explanation for this.

This factor was not investigated in the technology based observations comprising studies one, three and four. However, as stated under the discussion of factor '#I1 learner expectations' above, '#I2 response certitude' offers some explanation for the instances observed in which subjects ventured beyond KCR type elaborative feedback.

Factor #I3 Interpretation of task

The ten interviews comprising study two strongly support this hypothesis.

In their model of self-regulated learning, Butler and Winne (1995) suggest that learners' interpretation of a task is a factor directly affecting how learners engage with feedback. When faced with a task, self-regulating learners first use their knowledge and beliefs to construct an interpretation regarding what is required by a task. Based on this interpretation learners then set goals and determine the tactics and strategies required to achieve these goals. Upon receiving feedback learners adjust these goals, tactics and strategies as required. Based on this model, it was hypothesised that learners' interpretation of the task would affect whether they subsequently chose to view a feedback type.

As discussed in the report of study two (See page 159), the instructional roles played by multiple choice questions can be organised into two overarching categories. First, multiple choice questions serve a testing role, they provide a numeric score that provides learners with a measure of their learning. Second, multiple choice questions facilitate learning by enabling learners to identify gaps in their knowledge and, through the provision of feedback, assisting learners in closing these gaps. Determining the instructional goal of an activity plays a significant role in learners' interpretation of the task. When faced with an activity comprising multiple choice questions, this is largely defined by the emphasis placed on one or both of these testing and learning roles.

The subjects in study two generally associated multiple choice questions with testing their knowledge rather than as an opportunity to learn. Subjects interpreted the task as to first complete the activity and second, to do so with as high a score as possible. Learners' emphasis on the purpose of multiple choice questions as a testing tool was given as a reason not to view elaborative feedback types beyond KCR. Conversely, those who recognised the role played by multiple choice questions in promoting learning stated this to be a reason to view these feedback types.

Studies one, three and four did not investigate factor 'I3 interpretation of task' however this factor does offer a partial explanation for some of the behaviour observed in these studies. As mentioned above under 'Factor #I1 Learner expectations', these studies observed variation in behaviour across the subject cohort not completely explained by differences in subjects' performance. Factors 'I3 interpretation of task' and '#1 learner expectations', offer an explanation for this unexplained variation.

In addition, as discussed in section 'Why learners move beyond KCR threshold' in the report of study two (See page 173), factors 'I3 interpretation of task', '#I1 learner expectations' and '#I2 response certitude' offer some explanation for the instances observed in these studies in which subjects ventured beyond KCR type elaborative feedback.

Factor #I4 Level of expertise

Three studies identified in the literature review suggest that learners' ability in a subject may impact on the effectiveness of feedback (Roper 1977; Butler and Winne 1995; Handley 2002) and this was the basis for the hypothesis that learners' level of expertise affects whether they choose to view the feedback types.

The ten interviews in study two provided limited support for this hypothesis however study four found no significant effect for factor ‘#I4 level of expertise’ when this was included in the logistic regression analysis.

Factor #I5 Performance

All four studies in this research found performance to have a significant effect in determining the frequency to which subjects viewed the feedback types. Study one found subjects were more likely to view feedback types KCR, IB and EX following an incorrect response. Subjects three and four found performance to be a significant factor for all six feedback types. Having answered a question correctly subjects were significantly less likely to view feedback types KR, KCR, IB, EX, EI1 and EI2 and therefore more likely not to view feedback. The odds ratios for significant effects for factor ‘#I5 Performance’ in these three studies are shown in the following table. In each case, the effect was $p < 0.001$.

Table 78: Significant effects for Factor #I5 Performance

	No feedback	KR	KCR	IB	EX	EI1	EI2
Study one			0.002	0.38	0.22		
Study three	3.56	0.62	0.001	0.11	0.07	0.24	0.31
Study four	8.54	0.21	0.01	0.11	0.07	0.33	0.63

As described above in the reporting of hypothesis two, the variation in subjects’ performance across an activity accounted for much of the variation in subjects’ use of the feedback types.

Study two also provided strong support for the importance of performance in determining the frequency at which learners elect to view a feedback type. Nine of the ten subjects interviewed explicitly stated performance to be a factor determining whether they would proceed to view the elaborative feedback types after initially viewing KR feedback.

Factor #I6 Gender

Hanna (1976) found the effect of TBI feedback during multiple choice question based instruction on subsequent post tests was more significant for males than females. From this it was hypothesised that learners’ gender may affect the frequency at which they elect to view a feedback type.

Possible effects for factor ‘#I6 gender’ were investigated in the logistic regression analyses performed in studies one and three. No significant effects were found for any of the feedback types. Neither was gender stated to be a factor during the semi-structured interviews comprising study two. This research concludes that gender is not a significant factor.

Factor #I7 Error type

Three studies identified in the literature review found that the type of error made by learners when answering a question incorrectly affected learners' interactions with feedback. Dempsey and Litchfield (1993) found error type to be a determinant of time spent on TBI feedback. Tatsuoka (1984) and Birenbaum and Tatsuoka (1987) found that error type affected the extent to which different feedback types improved subsequent performance.

Error type was not included as a factor for analyses in studies one, three and four. Neither was error type referred to by the subjects interviewed in study two. However, a possible explanation is that this term describes a concept in the theoretical educational domain and is not likely to be familiar to the subjects.

No conclusions can be reached as to the significance of factor '#I7 error type'.

Factor #E1 Assessment

This research was unable to reach conclusions regarding the effect of factor '#E1 Assessment'. Morrison, Ross et al. (1995) found that subjects were 3.24 times more likely to view 'review screens' following an incorrect response if the items of technology based instruction being studied constituted part of the course assessment. This led to the hypothesis that learner preferences for TBI feedback types change depending on whether the task being performed relates to an assessment.

Strong support for this hypothesis was provided by study two. Nine out of the ten subjects interviewed stated assessment to be a factor that would affect whether they chose to view a feedback type. The interviews revealed three different scenarios which potentially affected whether they chose to view the feedback types. First, if the activity constituted an assessed piece of coursework, second if the activity constituted an assessed piece of work and was held in exam type conditions and third, if the activity itself is not assessed but a direct link exists between the activity and a subsequent assessment.

Study three investigated the latter scenario. Some student cohorts were assessed on the "Accounting Primer" and the "Pre-study Mathematics" courses via a formal exam upon their arrival at the host institution. In study three, factor '#E1 assessment' referred to this subsequent assessment of these cohorts. In other words, to whether students were subsequently assessed on the TBI material. Subjects in study two stated that this condition would result in them electing to view the elaborative feedback types more frequently.

In study three, assessment was found to have a significant effect on whether subjects elected to view feedback types KR, KCR and EI2. However, the result differed to that expected. Subjects were less likely to view these feedback types if the material was subsequently assessed and therefore more likely not to view any feedback. The odds ratios for the significant effects found in study three are shown in the following table.

Table 79: Significant effects for factor ‘#E1 assessment’

	No feedback	KR	KCR	IB	EX	EI1	EI2
Study three	1.55 ($p < .018$)	0.58 ($p < .004$)	0.73 ($p < .001$)				0.72 ($p = .0080$)

This result contradicts the hypothesis based on the results Morrison, Ross et al. (1995) although the meaning of assessment differs with this study and also the prediction from study two. No reasonable explanation emerged for this finding and the effect of assessment remains unclear.

Factor #E2 Feedback level

The models of feedback proposed by Kluger and DeNisi (1996) and Hattie and Timperley (2007) both make reference to the fact that feedback can be directed at different levels. For example, Kluger and DeNisi refer to three different levels of control a) task learning, b) task motivation and c) meta-task processes (See page 54).

Feedback level was not investigated in this research. An analysis of the TBI forming the focus of study one revealed that all existing feedback was at a single level, the task level.

This research, identifies the potential for ‘#E2 feedback level’ to be a factor affecting the frequency at which learners elect to view the feedback types but does not provide any empirical evidence that this may be the case.

Factor #E3 Time limit

The hypothesis that setting a time limit on an activity is a significant factor affecting whether learners chose to view a feedback type emerged from studies one and two.

In study one, subjects undertook the Ethics and Strategy exercise together in a computer laboratory. Care was taken in the preparation and design of this study to minimise the impact of a time limit. While the exercise generally takes between 25 and 30 minutes to complete, subjects were given 60 minutes. Subjects were told that they more than sufficient time and also that they were not required to answer all of the questions.

Despite these measures, seven of the ten respondents in study two stated the presence of a time limit to be a factor affecting whether they chose to view a feedback type. Subjects acknowledged that a surplus of time was available but stated that the mere presence of a time limit affected their behaviour.

It was not possible to investigate the effect of factor ‘#E3 time limit’ in studies three and four as all subjects studying the Accounting Primer and Pre-study Mathematics courses did so without a time limit and at a time and place of their choosing. It was therefore not possible to isolate time limit as a factor.

Factor #E4 Location

Study two also lead to the hypothesis that the location at which an activity is taken affects whether learners choose to view the feedback types. Seven of the ten respondents interviewed in study two stated that their location when taking the activity would affect their selection of feedback.

As discussed in the report of factor ‘#E3 time limit’ above, it was not possible to investigate the effect of factor ‘#E4 location’ in studies three and four as all subjects studying the Accounting Primer and Pre-study Mathematics courses did so without a time limit and at a time and place of their choosing.

Factor #E5 Question focus

In this research, factor ‘#E5 question focus’ refers to whether a question tests recall or application of the material studied. A hypothesis emerged from study two that question focus is a significant factor affecting whether learners choose to view a feedback type and support for this hypothesis was provided in studies three and four.

Three of the ten subjects interviewed in study two drew a distinction between questions that test the recall of course concepts and questions that require the application of these concepts in novel scenarios. These subjects stated that their choice of feedback would differ between the two questions types.

This factor was investigated further in studies three and four. All questions in the three items of TBI studied by subjects were marked as testing either recall or application of the material and question focus was included as a factor in the logistic regression analyses.

In study three, question focus was found to be significant factor in determining the frequency at which subjects elected to view feedback types KCR, IB, EX, EI1 and EI2 and also whether subjects chose not to view feedback. Subjects were significantly more likely to view these feedback types if a question tested application of the material studied. Study four found similar positive effects were found for feedback types KCR and EX however there were no significant effects for feedback types IB and EI1. The odds ratios for the significant effects found in studies three and four are summarised in table 80.

Table 80: Significant effects for #E5 Question focus

	No feedback	KR	KCR	IB	EX	EI1	EI2
Study three	0.78 (p=.0038)		1.57 (p<.001)	1.64 (p<.001)	2.28 (p<.001)	1.69 (p<.001)	1.48 (p<.001)
Study four			1.59 (p=.014)		1.80 (p<.001)		

From this we can conclude that learners are more likely to view KCR and EX type feedback if questions test application as opposed to recall of the course material. Question focus may also affect the frequency in which learners elect to view IB, EI1 and EI2 type feedback however this should be subject to further investigation.

Factor #E6 Subject matter

Related to factor ‘#E5 question focus’ is factor ‘#E6 subject matter’. The potential for this factor to affect whether learners choose to view a feedback type also emerged from study two.

Seven of the ten subjects interviewed in study two stated that they would be more likely to view the elaborative feedback type if the activity related to quantitative subjects such as accounting and mathematics. From this a hypothesis was formed that whether learners choose to view a feedback type is affected by whether a question quantitative or qualitative material. This is termed factor ‘#E6 subject matter’ in this research.

It was not possible to investigate ‘#E6 subject matter’ further in studies three and four. An analysis of the questions across the three items of TBI revealed that all of the questions in the Ethics and Strategy Exercise were qualitative while all of the questions in the Pre- Study Mathematics course were quantitative. It was therefore not possible to isolate ‘#E6 Subject matter’ from ‘#E5 question focus’ as a factor for investigation.

This research, provides some evidence that ‘#E6 subject matter’ is a factor affecting the frequency at which learners elect to view the feedback types but further investigation is required to confirm that this is the case.

5.1.7 Hypothesis 5: The majority of learners will revisit a previously viewed question

Data were gathered in studies one, three and four to investigate this hypothesis. The hypothesis was not supported. In study one, 86.30 per cent of students revisited at least one question over the course of the activity and on average, subjects revisited 4.17 questions. In addition, subjects elected to view an additional feedback item on 42.19 per cent of revisits. This led to the hypothesis that the majority of students will revisit a previously viewed question. These results, together with corresponding data from studies three and four are presented in table 81 on the following page.

Table 81: Summary data relating to subjects revisiting questions

Study one			
	Ethics and Strategy exercise		
Proportion of students who revisited at least one question	0.8630		
Mean number of questions revisited across the cohort	4.17		
Study three			
	Accounting Primer course section	Pre-study Mathematics course section	Mean across the two sections
Proportion of students who revisited at least one question	0.4356	0.4000	0.4285
Proportion of students who revisited question and viewed additional feedback	0.2228	0.1000	0.1984
Mean number of questions revisited across the cohort	2.10	0.74	1.83
Mean number of additional feedback items viewed across the cohort	0.56	0.20	0.49
Study four			
	Accounting Primer course section	Pre-study Mathematics course section	Mean across the two sections
Proportion of students who revisited at least one question	0.3947	0.3590	0.3769
Proportion of students who revisited question and viewed additional feedback	0.2105	0.21	0.2078
Mean number of questions revisited across the cohort	1.61	1.15	1.38
Mean number of additional feedback items viewed across the cohort	0.55	0.46	0.51

As can be seen from these data, while a significant proportion of subjects did revisit questions in studies three and four, the majority did not and therefore hypothesis five was not supported.

5.1.8 Hypothesis 6: Learners' use of feedback did not change significantly over the course of the activity

As discussed above, it was not possible to collect meaningful data for this hypothesis in studies three and four and no meaningful conclusions can be reached.

6. DISCUSSION

This research investigated learners' preferences for the format of TBI feedback and was guided by two questions:

1. What format of technology based instructional (TBI) feedback do learners elect to view if given a choice?
2. What are the factors that affect this choice?

The results of this research are now interpreted in relation to the theoretical framework and with reference to prior research and prior models of feedback. Results are now into eight themes and are discussed in turn. Following this, these eight themes are unified and a model of learner interaction with TBI feedback is proposed.

6.1 THE EIGHT THEMES

6.1.1 Theme 1: When provided, learners view feedback extensively.

The literature review identified a study by Pridemore and Klein (1995) which found that the provision of TBI feedback had a significant positive effect on learners' attitude. From this, a hypothesis was formed that learners will elect to view feedback more frequently than they do not. This hypothesis was supported. Study one found that subjects elected to view at least one feedback type for 85.19 per cent of the questions they answered. In study three this figure rose to 92.50 per cent and in study four this rose again to 94.67 per cent. Across all the studies, subjects viewed feedback for 92.72 per cent of the questions they answered. It is apparent from these studies that, if provided, learners will view feedback extensively.

A recurrent theme throughout this chapter is that constructivist views on the nature of the learner and the nature of learning accurately reflect the behaviour observed. The constructivist view of the learner, as adopted in this research, is described in the section 'Theoretical framework' (See page 65). Of pertinence here is that learners are viewed as active participants to a learning experience.

“Knowledge is constructed by learners as they attempt to make sense of their experiences. Learners are not empty vassals waiting to be filled but are active organisms seeking meaning” (Driscoll 2000: 376).

In this series of studies, learners were active participants. They made choices throughout and these choices varied from question to question depending on the factors identified in this research. Learners were active, they were seeking 'meaning' from their experiences and this behaviour was consistent with the constructivist view of the learner described above.

6.1.2 Theme 2: Learners view feedback in order to achieve closure on the task

In constructivist thought it is the identification and resolution of knowledge gaps that leads to learning. Ernst Von Glasersfeld (1989: 136) uses the term ‘perturbation’ to describe the state when a knowledge gap is identified and ‘equilibrium’ to describe the state at which closure is achieved regarding this perturbation. Learners actively seek to resolve perturbations. These notions of ‘perturbation’ and equilibrium’ provide a cohesive explanatory framework for the behaviour observed in this series of studies and also for subjects’ explanations of this behaviour.

Study two found that subjects’ reason for viewing feedback was to achieve closure on the task. In this research, ‘closure’ is defined as per Handley (2003),

“The perception, on the part of the learner that expectations – arising specifically from the task, and generically from learners’ expectations of what constitutes good feedback - are satisfactory met.”
(Handley 2003: 206)

When viewing the different feedback types, subjects pursued closure first by seeking affirmation as to the correctness of their response and second by seeking to resolve ‘knowledge gaps’, that is any gap between what a subject understood and what had needed to be understood in order to have responded to a question correctly. When seeking to resolve knowledge gaps, most subjects stated that their primary purpose was a need to understand and, to a lesser extent, a requirement to learn. Learners do not always seek to comprehensively resolve such gaps when they are identified. This is not always required for closure to be achieved.

These results are consistent with the constructivist notions of ‘perturbation’ and equilibrium’ and these concepts form one foundation of the model of learner interaction with TBI feedback proposed later in this chapter.

6.1.3 Theme 3: Subjects initially view KR type feedback and then view KCR type feedback if they have responded incorrectly

With regard to KCR and KR type feedback, a common pattern emerged. Subjects first elected to view KR type feedback following a response, they did so for 91.11% of the questions they answered. Following this, subjects generally elected to view KCR type feedback when they responded incorrectly, they did so 83.19% of the time.

Prior research had established that learners do use KR components to feedback if it is presented,

“There is ample evidence that, in instructional situations like the ones in these studies, learners use feedback to check the correctness of their answers.” (Bangert-Drowns, Kulik et al. 1991: 233)

This research suggests that learners will actively seek KR type feedback if it is not presented.

That subjects generally elected to view KCR type feedback following an incorrect response could be, at first glance, interpreted as TBI feedback playing the error correcting role proposed under the cognitivist paradigm. However, as discussed under theme one above, the pursuit of closure more accurately describes subjects’ behaviour. While

subjects did seek to correct errors, subjects did not always require an error to be corrected in order for closure to be achieved. As before, the constructivist notion of learning occurring through the identification and resolution of perturbations acts as an accurate overarching explanatory framework for the behaviour observed.

6.1.3.1 Note on the existence of a feedback threshold

The literature review identified the proposed existence of a feedback threshold, at the level of KCR type elaborative feedback, beyond which the study of feedback becomes either ineffective or inefficient (Phye 1979; Kulhavy, White et al. 1985; Dempsey and Litchfield 1993; Pridemore and Klein 1995).

Generally, elaborative feedback is more complex than verificational feedback and requires a significantly higher investment of time and resources to develop. In the educational context, a limiting resource is the educator's time and establishing priorities for this time may be required. As described earlier (See page 52), both Dempsey and Litchfield (1993) and Kulhavy et al. (1985) suggest that the time spent by educators in developing feedback beyond KCR is not justified by the subsequent effects on learning. These studies conclude that feedback types beyond KCR are inefficient using the following formula to calculate the efficiency of feedback.

$$\text{Feedback efficiency} = \frac{\text{correct retention percentage score}}{\text{square root of feedback study time}}$$

The premise of the present research is to provide learners with a choice of feedback types and then observe which types they elect to view. If the existence of a feedback threshold held for subjects in this research then we should expect to find that subjects elected to view feedback type KCR considerably more frequently than types IB, EX, EI1 and EI2.

In study one, subjects did elect to view feedback type KCR considerably more frequently than types IB, EX, EI1 and EI2. This provided support for the existence of a threshold. The results of study one indicated that subjects were aware of this threshold and only elected to view the elaborative feedback types beyond KCR on a small minority of occasions. However, such a threshold effect was less pronounced in studies three and four. In both studies the difference in mean α values did not give the appearance of a threshold. In particular, the mean α values for EX type feedback for the Pre-study Mathematics course was noticeably closer to the value for KCR type feedback. In other words subjects elected to view an explanation of how to answer a question correctly almost as often as they elected to view the correct answer.

An explanation for this lies in the nature of the questions in the Pre-study Mathematics course. The results of the investigation into hypothesis four found a significant effect for factor '#E5 question focus' on feedback type EX. Subjects were more likely to click on this feedback type when a question tested application rather than recall of the material. All questions in the Pre-study Mathematics course tested application of the material and this offers an explanation of the high mean α values value for feedback type EX. Certainly for the Pre-study Mathematics course, subjects' actions suggest they disagreed with the hypothesis that feedback beyond the level of KCR becomes either ineffective or inefficient.

The four studies identified in the literature review that refer to the existence of a threshold each involved different subject matters. The questions in Pridemore & Klein (1995) focused on the functioning of microscope, Dempsey and Litchfield (1993) focused on biology, Phye (1979) focused on psychology and questions in Kulhavy, White et al. (1985) tested the comprehension of written texts. While it is not possible to be certain, the impression given by the reports of these studies is that none of the questions in the TBI materials tested quantitative material nor tested application of the material studied.

This research suggests that the existence of a threshold is dependent on the subject matter of a test and whether questions test recall or application of the study material.

6.1.4 Theme 4: Subjects held a preference for elaborative feedback types beyond KCR that focus on the knowledge gap

The occasions on which subjects elected to view the feedback types beyond KR and KCR were less predictable. However, this research did identify factors that affect this choice and these are discussed under ‘theme 6’ below. In addition, this research concludes that the degree to which an elaborative feedback type focuses on an identified knowledge gap affects the frequency at which learners will elect to view that feedback type.

In this research it is proposed that the level of specificity on this knowledge gap varies across the feedback types adopted as follows:

- EX type provides an explanation as to why a learner’s response was incorrect. This feedback type is the most focused on the knowledge gap.
- IB type feedback presents a subject with the relevant section from the course material. This feedback type is sufficient to close the knowledge gap when a question recycles concepts verbatim from the course material. However, this feedback type may not be sufficient to close the knowledge gap when a subject is required to apply concepts from the materials to the question. In general, this feedback type is considered here to be less focused on the knowledge gap than EX type feedback.
- Extra-instructional or EI type feedback provide additional information relevant to a question.
 - EI1 type feedback provides further examples of the concepts introduced in a question or solutions to similar exercises. This information assists learners in closing the knowledge gap but is less specifically focused than either IB or EX type feedback.
 - EI2 type feedback provides explanations of terminology and rare words. This is not generally sufficient to close the knowledge gap and is generally the least focused feedback type.

The information in EI type feedback assists a learner to resolving a knowledge gap but, in itself, is generally not sufficient to close such gaps. This feedback type is thus less focused on a feedback gap than types EX and IB.

Should this hypothesis hold then we should expect mean α values for KCR type feedback to be the highest followed by those for EX then IB and then the extra-instructional feedback types EI1 and EI2. This hypothesis generally held but there was one exception. The mean α value for IB type feedback for subjects studying the Accounting Primer course was higher than the value for EX type feedback. However, a possible explanation exists. As described in the reports of studies three and four above (See pages 203, 225), 70.82 per cent of the questions responded to by the subjects in the Accounting Primer course tested recall of the course material and it was argued that, for these questions, the presentation of the relevant course material via IB type feedback provides an explanation of the correct answer equally as efficiently as the explanations provided by EX type feedback. In addition, by re-viewing the original material, the learner can readily identify the nature of any misunderstandings that arose during their study of the material. Thus, for the Accounting Primer course, IB type feedback can often be considered to focus more directly on any knowledge gaps than EX type feedback.

Subjects' preference for elaborative feedback types beyond KCR that focus on knowledge gap is explained by their focus on the pursuit of closure. When describing their behaviour, subjects in study two rarely referred to goals at the wider course or activity level, their focus was at the question level and in particular in achieving closure on a question. Equivalently here, this can be considered as a focus on resolving perturbations following the viewing of KR or KCR type feedback.

In addition to constructivist views on the nature of the learner and the nature of learning, the theoretical framework in this research incorporates cognitive load theory and schema theory. The former theory, discussed in the section 'Theoretical framework' (See page 65), offers some explanation through its focus on the mechanisms of working memory. Proponents of cognitive load theory point to the limitations of this memory,

“Because working memory is most commonly used to process information in the sense of organizing, contrasting, comparing, or working on that information in some manner, humans are probably only able to deal with two or three items of information simultaneously when required to process rather than merely hold information.” (Sweller, van Merriënboer et al. 1998: 252)

This notion of limits to working memory is referred to on a number of occasions during this chapter. Here, it offers one explanation for learners' focus on seeking closure, or resolving perturbations, at the question level when feedback is presented immediately following a response.

6.1.5 Theme 5: Subjects elect to view the minimum load of TBI feedback required to achieve closure

Study two found that subjects sought to view the minimum load of TBI feedback required to achieve closure and offered three reasons for this. First, subjects stated a wish to be time efficient. Subjects considered time to be a precious resource and were keen to minimize the time spent on feedback. They wanted to access the information they required from feedback as fast as possible. Second, some subjects stated that a high feedback load dampened their sense of progress through an activity. These two issues are explained by subjects' interpretation of the task. The subjects in study two generally exhibited performance orientated behaviour. They associated multiple choice

questions with testing their knowledge rather than as an opportunity to learn. Subjects generally interpreted the task faced as to first complete the activity and second, to do so with as high a score as possible. This is discussed in more detail below under 'Factors at the learner level'.

The third reason proffered for subjects wishing to view a minimal TBI feedback load was that high volumes of feedback are difficult to absorb. Again, cognitive load theory, provides an theoretical explanation for this view highlighting the limit on the number of concepts that can be held and processed within working memory if learning is to occur.

This behaviour also reflects subjects' strong focus on achieving closure or, to be consistent with the constructivist framework adopted in this chapter, resolving perturbations.

6.1.6 Theme 6: Factors at the question, learner and environmental level affect whether learners choose to view the TBI feedback types

The variation occurring in this research can be considered in two dimensions. First, individual learners elect to view different feedback types for different questions in a test. Second, some learners will generally elect to view certain feedback types more often than others. The factors that cause these variations are now discussed in terms of:

- Factors at the question level
- Factors at the learner level
- Factors at the environmental level

6.1.6.1 Factors at the question level

This research established that a number of factors relating to learners' engagement with an individual question affected whether they chose to view the TBI feedback types.

The four studies suggest that factor '#I5 performance', whether a learner responds correctly to a question, to be one such factor. That learners will elect to view the TBI feedback types more frequently following an incorrect response appeals to common sense and personal experience and this research suggests that this is the case. Study one found subjects were more likely to view feedback types KCR, IB and EX following an incorrect response. Studies three and four found performance to be a significant factor for all six feedback types. Having answered a question correctly subjects were significantly less likely to view feedback types KR, KCR, IB, EX, EI1 and EI2 and therefore more likely to not to view feedback.

Subjects did, however, elect to view the elaborative feedback types following an incorrect response and the subjects in study two stated a rationale for this, they would do so having guessed the correct answer. This has potentially

valuable implications for designers of instructional materials and this is discussed further in ‘Appendix two: Recommendations for practitioners’.

Study two established that factor ‘#I2 response certitude’, the degree of certainty learners hold regarding the correctness of their response, to be a factor affecting whether learners elect to view a feedback type. Seven of the ten subjects stated response certitude to be a significant factor in choosing to move beyond KR and KCR type feedback and view the IB, EX, E1 and E2 elaborative feedback types. All seven subjects stated they elected to view the additional feedback when they held a high level of certainty that their answer had been correct. The concept of response certitude was initially explored by Kulhavy and Stock (1989) whose resulting model predicts that learners will spend more time on feedback following an incorrect response if their level of response certitude is high. Study two revealed an additional insight into why this is the case, this being that learners give attention to more information in the feedback message when holding a high degree of response certitude. This is discussed further in the section ‘Implications of the results’ (See page 265).

In addition to factors ‘#I5 performance’ and ‘#I2 response certitude’, the literature review identified a further factor relating to learners’ engagement, this being factor ‘#I7 error type’, the type of error made when responding to a question. However, this factor was not investigated further in studies three and four.

A further factor identified as being significant relates not to learners’ engagement with a question but the nature of the question itself. Study two identified factor ‘#E5 question focus’, whether a question tests recall or application of the material, to be a potential factor. In studies three and four this factor was found to affect the frequency at which learners elect to view KCR, EX and possibly IB, EI1 and EI2 type feedback. This has implications both for interpreting prior research and also for educators involved in the design of TBI materials. This is discussed in detail in the sections ‘Implications of the results’ (See page 265) and ‘Appendix two: Recommendations for practitioners’. A related factor ‘#E6 subject matter’, whether a question tests subject matter that is qualitative or quantitative, was also identified as being potentially significant however, it was not possible to investigate this further in this research.

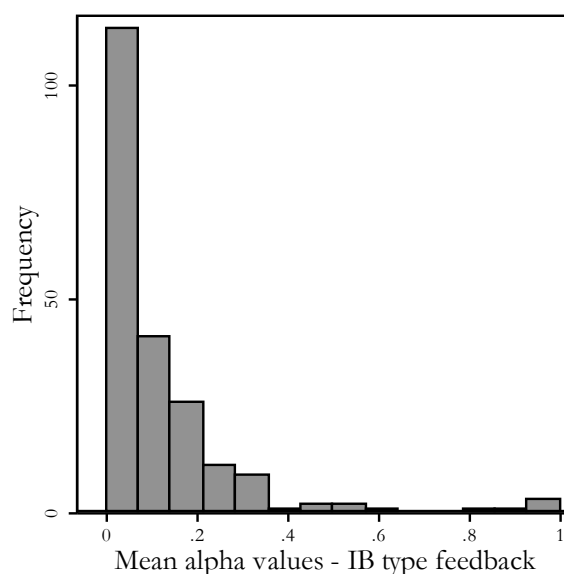
A theme throughout this chapter is that the constructivist notion of learning occurring through the identification and resolution of perturbations. The effects of factors ‘#I5 performance’, ‘#I2 response certitude’ and ‘#E5 question focus’ can be explained as affecting the appearance and nature of a perturbation. The mechanisms through which this occurs is discussed in more detail in the model of learner interaction with TBI feedback proposed later in this chapter.

6.1.6.2 Factors at the learner level

In addition to factors at the question level, this research suggests that some learners have a general tendency to view the elaborative feedback types more frequently than other learners. When investigating the distribution of α values for the elaborative feedback types across the subject cohort, a common pattern was established. A significant proportion of the subjects did not elect to view these feedback types. The majority of subjects viewed them rarely while a minority of students elected to view these feedback types across a range of frequencies. Thus, when plotted on a graph, the distributions of α values across for all cohorts in this series of studies contained a marked positive

skew for the KCR, IB, EX, EI1 and EI2 elaborative feedback types. This is illustrated in the following histogram which plots the α values for IB type feedback relating to the subjects in study four.

Figure 56: α values for IB type feedback for subjects in study four



Factor ‘#I5 performance’ was the dominant factor and differences in subjects’ overall performance accounted for the majority of this variation. However performance did not account for all the variation observed. The literature review identified factors ‘#I6 gender’ and ‘#I4 level of expertise’ to be factors that potentially explain the remaining variation however studies one and four established that these factors not to be significant. However study two does provide some explanation by suggesting that factors ‘#I1 learner expectations’, through subjects’ style of engagement and sense of enquiry, and ‘#I3 Interpretation of task’ account for some of the remaining variation in subjects’ behaviour.

These factors are related. Learners’ interpretation of the task is one determinant of their subsequent style of engagement via the setting of goals. Butler and Winne’s (1995) model of feedback and self-regulated learning provides a theoretical link between learners’ interpretation of a task, their setting of goals and their subsequent engagement with the task. This model states that, when faced with a task, self-regulating learners first use their knowledge and beliefs to construct an interpretation of what is required. Then, based on this interpretation, learners subsequently set goals and determine the tactics and strategies required to achieve these goals.

A number of studies have investigated the effect of the goals learners’ set themselves when engaging with instructional materials (Ivancevich and McMahon 1982; Grant and Dweck 2003). Much of this research places learners’ goals into two overarching categories. First, learning goals in which learners aim to acquire new knowledge or skills. Second, performance goals in which learners aim to validate their ability or avoid demonstrating a lack of ability. Consistent with much prior research, Grant and Dweck (2003) found positive effects for learning goals on both learners’ motivation and performance. The same study classifies performance goals into outcome goals (wanting to do well), ability-linked goals (wanting to validate perceived ability) and

normative performance goals (comparing one's ability to that of peers). Grant and Dweck found that both ability-linked and normative goals had a significant effect on performance and that ability-linked goals had a significant effect on motivation.

In study two, subjects' interpretation of the task fell into two categories closely resembling these overarching categories of goals. Subjects generally interpreted the task as either a performance exercise or an opportunity to acquire new knowledge, with the majority interpreting solely as the former. It can be assumed that this interpretation resulted in learners being primarily focused on either performance goals or outcome goals and that this in turn determined their subsequent style of engagement. Thus, Grant and Dweck (2003) and Butler and Winne (1995) provide a theoretical framework linking interpretation of the task and a subsequent style of engagement.

Learners' sense of inquiry, or level of interest, also affects their style of engagement. In Keller's ARCS' model of motivation in instruction (1987: 4) the level of interest is one component of perceived relevance, the 'R' in ARCS. Relevance, together with Attention, Confidence and Satisfaction are given as the four conditions that need to be met for learners to become and remain motivated when engaging with instructional materials. Thus subjects' sense of inquiry affects their motivation and this in turn affects subjects' style of engagement.

There has been a great deal of research investigating learners' level of interest, motivation, goal setting and style of engagement (Ivancevich and McMahon 1982; Keller 1987; Grant and Dweck 2003; Handley 2003). The resulting theoretical landscape is multifaceted however students exhibiting behaviour that veers towards either performance or learning orientated behaviour is a concept that is present throughout. These two behaviours also emerged as strong themes in the series of interviews comprising study two and provide one explanation for the variation in the α values not explained by variation in performance.

6.1.6.3 Factors at the environmental level

Study two suggested that factors '#E3 time limit' and '#E4 location' affect the frequency at which learners elect to view a TBI feedback type. Seven of the ten respondents stated both of these factors would affect whether they chose to view a feedback type.

It was not possible to investigate the effects of these factors further in studies three and four as all subjects studying the Accounting Primer and Pre-study Mathematics courses did so without a time limit and at a time and place of their choosing. It was therefore not possible to isolate the factors for investigation. However, these studies do hint at these factors being significant. In study three, the largest effect among all factors included in the logistic regression related to the item of TBI being studied. The odds ratios across the feedback types ranged from 0.11 to 10.19. Subjects were far more likely to view feedback types IB, EX, EI1 and EI2 when studying the Accounting Primer and Pre-study Mathematics tests than when studying the Ethics and Strategy test.

Data from the Ethics and Strategy test were not included in study four. While test type remained a significant factor, the effect was markedly smaller and only applied to feedback types KR and KCR and the no feedback

condition. Subjects studying the Pre-Study Mathematics course were less likely to view KR type feedback (OR=0.21) and KCR type feedback and therefore more likely not to view feedback (OR=7.48). There was no effect of test for the remaining feedback types.

Thus it can be concluded that the large effect for the item of TBI being studied in study three was due to the presence of the Ethics & Strategy exercise. One explanation is that this effect was due to the different questions and feedback in this exercise however the conditions in which the subjects took the activity were also different. While subjects studying the Accounting Primer and Pre-study Mathematics courses did so at a time and place of their choosing, subjects took the Ethics & Strategy exercise together in a computer laboratory and with a time limit. Given the findings of study two it is plausible that these changes in conditions also contributed to the marked effect for the item of TBI being studied found in study three.

This evidence is insufficient to draw firm conclusions. However, if these factors are found to be significant then this would provide valuable information to educators designing blended or distance learning programmes. This is discussed further in the section 'Suggestions for future research' (See page 272).

6.1.7 Theme 7: Learners prefer immediate as opposed to delayed feedback

There is an on-going debate in the literature as to whether TBI feedback should be provided immediately after a response to a question or following a delay. Generally, studies provide support for the provision of immediate feedback (Kulik and Kulik 1988; Dempsey, Driscoll et al. 1993; Dihoff, Brosvic et al. 2003). However a significant number of studies found performance benefits in delaying the provision of feedback (Dempsey, Driscoll et al. 1993; Morrison, Ross et al. 1995; Butler, Karpicke et al. 2007). This research investigated learners' preferences for the timing of feedback. In other words, do learners elect to view immediate or delayed feedback when given the choice?

The results were conclusive. Prior to starting the exercise, subjects were given the choice of viewing feedback after responding to a question (immediate feedback) or viewing feedback at the end of the activity (delayed feedback). At this initial stage 28.17 per cent of subjects elected to view delayed feedback. After starting the activity subjects were able to change this setting and 65.57 per cent of subjects who had initially selected delayed feedback did so. As a result, 95.22 per cent of subjects ended their activity electing to view feedback after responding to a question with just 4.78 per cent of subjects electing to view feedback after responding to all the questions.

Thus subjects expressed a clear preference for immediate feedback. The subjects interviewed in study two provided an explanation for this, subjects considered immediate feedback to be more efficient. Subjects stated that receiving feedback immediately after responding to a question was more efficient because they would already be holding the detail of the question in mind. Receiving feedback at the end of the exercise would require them to re-study a question in order to recall the thought behind their response and therefore required more effort.

The theoretical framework adopted in this research provides a theoretical explanation for this behaviour, specifically schema theory. As discussed in the section 'Theoretical framework' (See page 65), schema theory

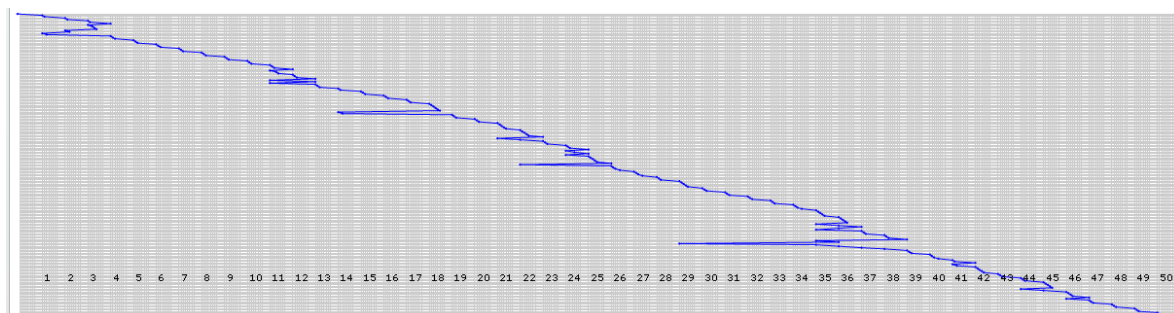
emphasises the distinction between long term and working memory. Knowledge is stored in long term memory in the form of schemas. In order for new knowledge to be acquired, relevant schemas need to be activated and loaded into working memory (Derry 1996: 167). This activation, refinement and generation of schemas form a significant part of the cognitive mechanisms that result in learning. In the current context, it is proposed that the activation of relevant schemas occurs at the point of responding to a question and thus when feedback is presented immediately, the relevant schemas are already activated.

In their review of prior research, Hattie and Timperley (2007: 98) point to some studies proposing that the effects of delayed feedback may be related to the nature of the task and also to the difficulty of the questions presented. As such, caution should be applied when generalizing this result. However, it is clear that subjects in this research preferred immediate feedback.

6.1.8 Theme 8: Learners will revisit questions if allowed

In this series of studies a significant proportion of subjects chose to revisit questions during a study session. As detailed in the section ‘The Development of the Software Tool’, the online system tracks subjects’ pathways through the multiple choice questions presented. This is shown in figure 57 below which shows the pathway for one subject. It is therefore possible to establish whether a subject revisited questions and if they did whether they chose to view an additional item of feedback.

Figure 57: An example of a subject’ pathway



In study one, 86.30 per cent of students revisited at least one question over the course of the activity. This led to the hypothesis that the majority of students will revisit a previously viewed question. This hypothesis was not supported in studies three and four in which respectively only 42.85 per cent and 37.69 per cent of subjects revisited questions.

This can be explained by the difference in the length of the activities. The test in the Ethics and Strategy exercise from study one consists of 50 questions. Tests in Accounting Primer and Pre-study Mathematics courses are divided into shorter tests containing between four and seven questions. The figures of 42.85 per cent and 37.69 per cent refer to revisits made when students studied smaller tests comprising six questions. Given the relatively short length of these tests, a possible explanation for the reduction in revisits compared to study one is that subjects were more likely to hold memory of prior questions throughout the test. While hypothesis five was not supported in studies three and four, the proportion of subjects who did revisit questions is notable.

The ability to revisit questions is considered here a form of feedback. In this research, subjects' held a strong focus on achieving closure or, equivalently, resolving perturbations at the question level. It is assumed here that when subjects' elect to revisit a previously viewed question, they do so as part of this process. For example, revisiting a similar question may help to resolve a perturbation relating to the question at hand.

6.2 A MODEL OF LEARNER INTERACTION WITH TBI FEEDBACK

These eight themes are now unified and integrated with aspects of prior theoretical models to produce a model of the process through which learners engage with the TBI feedback types in the taxonomy of TBI feedback.

As discussed the section 'Theoretical framework' (See page 65), the overarching framework through which the results of this research are interpreted is constructivism. Constructivist thought informs assumptions regarding the nature of reality and knowledge, the nature of the learner and the nature of learning. Two additional frameworks from the field of cognitive psychology are adopted to add detail to the latter, schema theory and cognitive load theory.

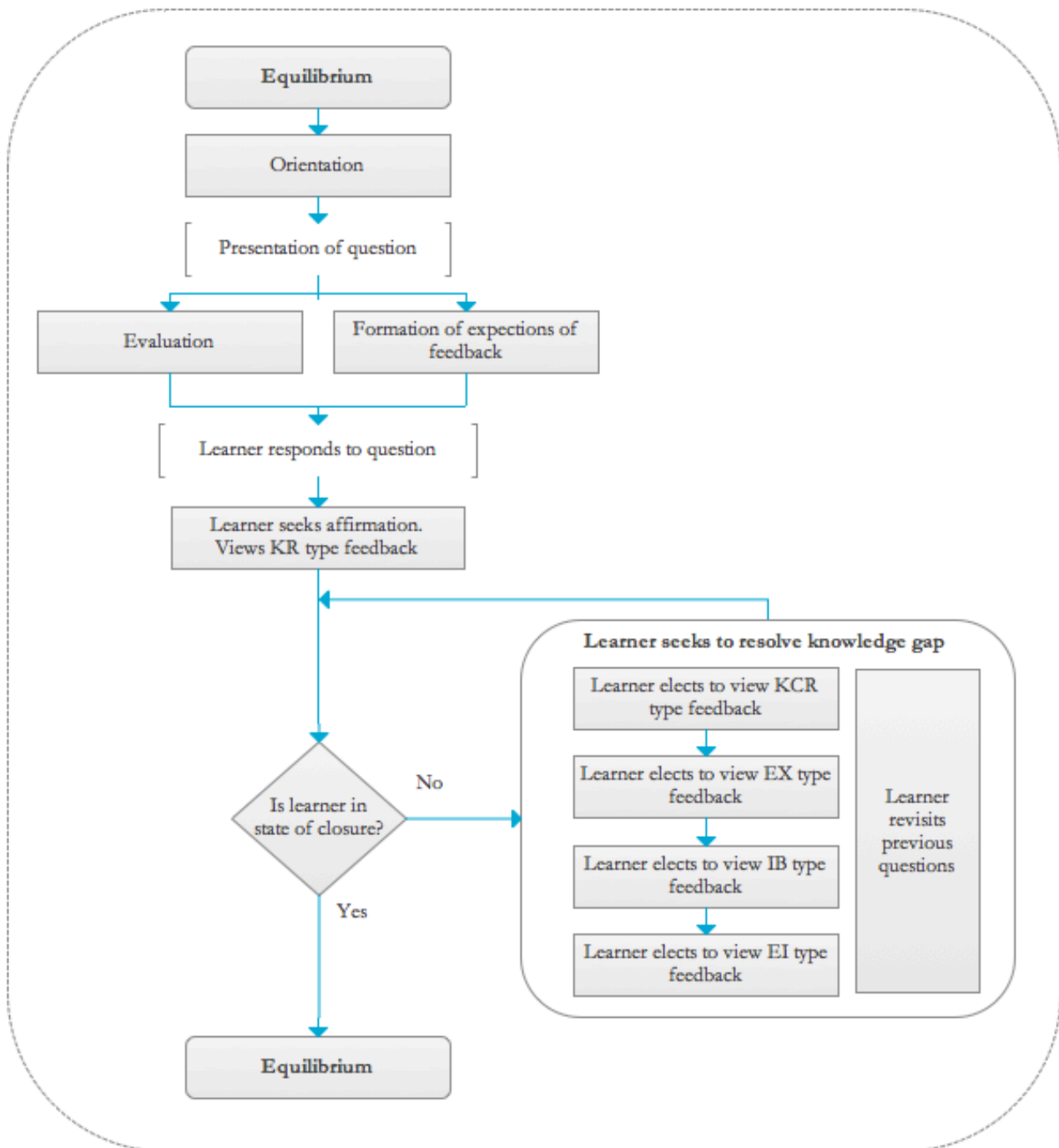
The model that follows draws heavily on these frameworks and, in addition, adopts aspects of the structure of the five-stage model proposed by Bangert-Drowns, Kulik et al. (1991) and the concept of closure proposed in Handley's (2003) model "Learners' expectations of, and responses to TBI feedback".

Central to this model is the constructivist notion of learning occurring through the identification and resolution of 'perturbations' arising from 'knowledge gaps'. This is introduced in the section 'Constructivism' under "theoretical framework' (See page 65) and is discussed further here. In the context of multiple choice questions, this knowledge gap is the gap between what was understood and what had been required to be understood in order to have answered a question correctly. Ernst Von Glasersfeld (Von Glasersfeld 1989: 136) proposes that learners begin a learning experience in a state of 'equilibrium' and uses the term 'perturbation' to describe the state when a knowledge gap is identified. Equilibrium is reached once more when the perturbation is resolved. This notion of 'equilibrium' forms the starting point of the model below and is defined by Von Glasersfeld (1989) as follows,

"Equilibrium refers to a state in which an epistemic agent's cognitive structures have yielded and continue to yield expected results, without bringing to the surface conceptual conflicts or contradictions. In neither case is equilibrium necessarily a static affair, like the equilibrium of a balance beam, but it can be and often is dynamic, as the equilibrium maintained by a cyclist." (Von Glasersfeld 1989: 126)

These theoretical frameworks together with the six themes described in the previous section form the basis for the model of the process through which learners engage with TBI feedback types now described. The model is first presented diagrammatically in figure 58 below.

Figure 58: A model of the learner engagement with the TBI feedback types.



Learners are assumed to begin an activity in a state of equilibrium with regard to the content of the activity. This is followed by a process of orientation. This refers to learners' activity immediately prior to being presented with a question. At this stage learners will have already begun forming expectations that will affect whether they subsequently choose to view the feedback types. They will have noticed environmental factors such as the time and location of the activity, will have interpreted elements of the task and will have formed a sense of inquiry. Thus the expectations that affect whether learners choose to view the feedback types are partially formed during this orientation however they are not fixed at this stage. They will be confirmed or revised as learners evaluate the question and the feedback types they choose to view.

Following the orientation, learners are presented with a question and this is followed by a process of evaluation. Kulhavy and Stock (1989) describe this process through the framework of servocontrol theory in which the learner matches the stimulus provided by a question to existing knowledge referents via a mechanism called a comparator. The model proposed by Butler and Winne (1995) describes this process as a cognitive system in which learners set goals and then determine tactics and strategies that produce 'products' on which learners' performance is subsequently evaluated. In the model of Bangert-Drowns, Kulik et al. (1991) this refers to the activation of search and retrieval strategies. In this research 'evaluation' refers to the process by which learners determine their response, this model does not attempt to provide detail on this process beyond the fact that this involves the activation of relevant knowledge schemas into working memory.

In parallel to this process of evaluation, learners form expectations regarding the feedback they will receive following their response. These expectations comprise those formed during the period of orientation plus expectations arising specifically from the question. Environmental factors, interpretation of the task and sense of enquiry affect these latter expectations as before. However, the most significant factor is a learner's degree of response certitude, that is the degree to which they believe that their response is correct. Kulhavy and Stock (1989: 293) describe how learners derive their degree of response certitude through reference to existing knowledge such as prior performance on similar tasks together with their level of knowledge in the subject being studied. In addition to forming expectations, learners may already experience a challenge to their existing knowledge framework, a perturbation, at this stage.

After evaluating the question and forming expectations regarding feedback, learners respond to the question and then actively pursue affirmation as to the correctness of their response by electing to view KR type feedback. Given the choice, learners will elect to view feedback immediately following their response to a question. Immediate feedback is the efficient choice as relevant schemas are already activated within working memory. KR type feedback serves to identify or confirm the presence of a perturbation or to confirm a state of cognitive equilibrium. Performance and response certitude affect the presence and strength of any perturbation.

Should a perturbation not be present then closure is achieved at this point and the process ends. If a perturbation is experienced then the learner will actively work to resolve this perturbation and generally proceeds to view the elaborative feedback types. These feedback types assist learners in achieving the accommodation of new knowledge. However, comprehensive resolution of a perturbation is not required in order for closure to be achieved. Environmental factors together with learners' expectations, interpretation of the task and response certitude affect the degree to which a perturbation needs to be resolved. Learners generally interpret objective questions as a tool to test their knowledge rather than an opportunity to learn. In addition, they generally elect to view the minimum load of TBI feedback required to achieve closure. This generally results in learners first electing to view KCR type feedback. If subjects elect to view additional elaborative feedback types beyond KCR then they prioritize viewing those feedback types most strongly focused on the knowledge gap identified by a question. They first view EX type feedback, then IB type feedback and finally EI type feedback. If a question tests application of the material, then learners are more likely to move beyond KCR type feedback at this stage. Learners also seek alternative sources of

information at this stage such as previous questions or materials outside of the activity. Learners pursue TBI feedback until closure has been attained

This completes the discussion of the results.

7. CONCLUSIONS

This chapter will first focus on the implications of the results and analysis reported in the previous chapter. A critique of the design and conduct of the research is then presented together with a discussion on the limitations of the research. The closing section of this chapter contains a number of suggestions for future research.

7.1 IMPLICATIONS OF THE RESULTS

The rationale for this research started with the premise that while feedback has a powerful effect on learning (Hattie and Timperley 2007: 83), prior research aiming to establish the format of TBI feedback that should be provided to learners is inconclusive. Such prior research focused on the relationship between the format of TBI feedback and scores achieved in subsequent performance tests. The inconsistent pattern of results led some to call for research placing a greater focus on the learner (Mory 2003: 771, Handley, Price, et al. 2011: 548).

The present research provides a greater level of understanding of learners' engagement with TBI feedback by identifying learners' preferences for the format of such feedback. In the previous chapter, these preferences were organised into eight themes which were then unified and integrated with prior theoretical models to produce a model of learner interaction with TBI feedback. This model identifies a process through which learners interact with different TBI feedback attributes and is a novel contribution to the body of knowledge relating to learners' engagement with TBI feedback.

This model has implications for research aiming to establish the effectiveness of different formats of TBI feedback, for existing theoretical models and for practice. These implications are now discussed.

7.1.1.1 Implications for research aiming to establish the effectiveness of different formats of TBI feedback

As discussed in the literature review (See page 46), research into TBI feedback is yet to identify the most effective format of such feedback. The model proposed in the present research has a number of implications for future research seeking to address this issue.

The model contends that factors at the question level, learner level and environmental level affect learners' interactions with TBI feedback. Further research is required to establish that these interactions affect the effectiveness of TBI feedback however it can be hypothesized that the effectiveness of a TBI feedback attribute depends on whether it is viewed by a learner and, if so, the degree of attention applied. Should this be the case then the implications are as follows.

The literature review noted substantial variability in the design of prior research (See page 47) however this research suggests that a number of factors relating to the design of a research study affect learners' subsequent interaction with TBI feedback. These include '#E5 question focus', '#E3 time limit', '#E4 location' and '#E6 subject matter'. A lack of control for these factors may have confounded prior attempts to generalise results.

It was also noted in the literature review (See page 47) that prior research did not control for factors relating to a learner and a learner's interaction with TBI feedback. This research suggests that factors '#I1 learner expectations', '#I3 Interpretation of task', '#I5 performance' and '#I2 response certitude', including learners' 'guessing' a correct response, all affect learners' interaction with TBI feedback. Again, the implications of this research are that the presence of these factors may explain the inconsistency found in prior research and that future studies should include appropriate controls.

Possible approaches for research aiming to identify the most effective format of TBI feedback are discussed in more detail in a following section 'Suggestions for future research' (See page 272).

7.1.1.2 Implications for existing theoretical models

The results of this research provide support for some existing theoretical models but challenge others. These implications are now explained and discussed.

Consistency with the view of feedback within the three dominant theoretical perspectives

The literature review referred to three dominant theoretical perspectives within educational research; behaviourism, cognitivism and constructivism and discussed how each perspective holds a different view on the role of educational feedback (See page 29). The degree to which the behaviour observed in this research is consistent with these views is now discussed.

Supporters of the behaviourist perspective see the role of feedback as a 'motivator' and a 'reinforcer' of correct behaviour (Mory 1992: 6), the learning and teaching equivalent of the food pellet used in operant conditioning. In this research the frequency at which subjects elected to view feedback types beyond KR indicates they wanted to know not just whether their response was correct but, if they were not, they also wanted to know why their response was incorrect. The fact that subjects elected to view such elaborative feedback, particularly for questions answered incorrectly, suggests that the behaviourist view is not able to explain all aspects of learners' interactions with feedback. It may be the case, in this series of studies, that responding to a question correctly reinforced existing correct knowledge however the design of this research was such that this information was not captured. However, it is the case that no reference was made by the subjects in study two to such reinforcing behaviour.

Cognitivists see feedback primarily as an error correction mechanism (Mory 2003: 747). In this research, the fact that performance, whether a learner answers a question correctly or incorrectly, significantly affected whether subjects chose to view feedback initially suggested that learners agree with this role. However, study two established that the achievement of closure does not always require a comprehensive resolution of the gap between what was understood by a learner and what had been required to be understood in order to have responded

correctly. Thus, from the learners' perspective, the cognitivist view only partially explains the role of feedback. This has implications for the accuracy of the models proposed by Kulhavy and Stock (1989) and Bangert-Drowns, Kulik et al. (1991) which are both premised upon the cognitivist paradigm.

The constructivist view is that feedback acts as a guide, assisting learners in the process of seeking meaning during their engagement with instructional tasks (Handley 2003: 28). In the present research, this view provides the most accurate framework to describe the behaviour observed. Subjects did appear to use feedback as a guide, referring to the different feedback types as they felt necessary. Subjects' behaviour as they interacted with the different types also reflected the constructivist model of learning occurring through the occurrence of knowledge gaps ('perturbations') and the resulting path to closure (return to 'equilibrium').

The existence of a 'feedback threshold' does not hold for all items of TBI

The literature review identified the proposed existence of a feedback threshold, at the level of KCR type elaborative feedback, beyond which the study of feedback becomes either ineffective or inefficient (Phye 1979; Kulhavy, White et al. 1985; Dempsey and Litchfield 1993; Pridemore and Klein 1995).

As discussed above in the section 'Note on the existence of a feedback threshold' (See page 251), this research suggests that the existence of a threshold is dependent on factor '#E6 subject matter', the subject matter of a test, and on factor '#E5 question focus', whether questions test recall or application of the study material.

Contribution to Kulhavy and Stock: The Place of Response Certitude

Study two provided support for the model proposed by Kulhavy and Stock (1989) relating to the role of response certitude, that is the degree of certainty learners hold regarding the correctness of their response. This model makes two main predictions. First, that response certitude affects the length of time learners spend processing a feedback message. Second, response certitude determines the degree to which feedback affects learners' performance in subsequent tests.

The model proposed by Kulhavy and Stock (1989) predicts that learners' level of 'response certitude' will determine the length of time they will spend on feedback. The model predicts that when subjects hold a high degree of response certitude but respond incorrectly then they will spend a longer time on feedback due to the level of information processing required,

"In this instance, the successful feedback cycle must eliminate the error response, substitute the correct, and strengthen the new response to the extent necessary for it to appear on the post-test. These multiple activities require considerable effort on the part of the learner. Because a high certitude error is the point of maximum discrepancy, it is also the point at which the model predicts the largest feedback processing times. The model also predicts that 'large scale' elaborative feedback will have the greatest effect when response certitude is low." (Kulhavy and Stock, 1989: 297)

The present research reveals a further level of detail on why learners spend more time on feedback following an incorrect response. This being that users give attention to more information in the feedback message when holding a high degree of response certitude.

Kulhavy and Stock's model also predicts that 'large scale' elaborative feedback will have the greatest effect when response certitude is low. In addition to supporting this prediction the present research provides a novel explanation, subjects stated that having guessed at a question they would be more likely to view the types of elaborative feedback beyond KCR.

Contribution to the debate regarding the timing of feedback

This thesis has made frequent reference to the on-going debate in the literature as to whether TBI feedback should be provided immediately after a response to a question or following a delay. Some studies provide support for the provision of immediate feedback (Kulik and Kulik 1988; Dempsey, Driscoll et al. 1993; Dihoff, Brosvic et al. 2003). Others find performance benefits in delaying the provision of feedback (Dempsey, Driscoll et al. 1993; Morrison, Ross et al. 1995; Butler, Karpicke et al. 2007).

This research adds a new perspective to this debate by asking the question, do learners elect to view immediate or delayed feedback when given the choice? The results were unambiguous; subjects demonstrated a clear preference for immediate rather than delayed feedback.

7.1.1.3 Implications for practice

Despite making no claims as to the effectiveness of different feedback formats, the results of this research do imply a number of recommendations for good practice relating to the format and presentation of TBI feedback. These are provided in the appendices to this thesis in the section "Appendix two: Recommendations for practitioners".

In addition, the results of this research will enable researchers and practitioners to identify any differences between the behaviour that learners prefer to adopt and any behaviours that are found to lead to more effective learning. As Laurillard (2002) states,

"An insight into the student's view of the learning process would give us some basis for deciding on a teaching strategy." (Laurillard 2002: 41)

7.2 CRITIQUE OF THE DESIGN AND CONDUCT OF THE RESEARCH

As detailed in the section "The three items of Technology Based Instruction (TBI)" (See page 102), a key challenge faced while conducting this research was the removal of the Ethics and Strategy exercise from the curriculum at the host institution following study one. In the design adopted at the outset of the research, learners' preferences for the format of feedback within this single item of TBI were to be investigated under a range of different conditions.

The need to focus on alternative items of TBI in studies three and four increased the complexity in the research design.

Further challenges relate to the decision to conduct exploratory research within an authentic learning environment. This was determined to be the necessary approach for the reasons detailed in the section 'Research methods' (See page 84). However this did lead to a complex and 'messy' research process. A successful exploratory approach will frequently lead to variables and hypotheses emerging after the initial planning stage and in the present research this led to challenges regarding the design of the technology based observations. The adoption of an authentic learning environment compromised the degree of control held over the research environment throughout the process. Specific issues relating to these two challenges are now discussed.

7.2.1 The presence of a delay between study one and study two

In the semi-structured interviews comprising study two, subjects were asked to recall their experiences of the activity studied in study one. As discussed in the report of study two (See page 148), it was not possible to conduct these interviews until one month following study one due to subjects' schedules and this resulted in a compromise to the design of the study.

A dilemma was faced. If subjects had been asked to repeat the activity then their responses may have referred to this repeat of the activity and not their experience in study one. However, if subjects did not look at the activity prior to being interviewed then their responses may have been compromised by an inability to recall the detail of the activity. As a compromise, subjects were told that they would be interviewed on their experiences in study one and were asked to familiarise themselves with the activity prior to being interviewed.

7.2.2 Potential presence of selection bias in study two

As also discussed in the report of study two, the selection of subjects in study two contained an element of self-selection and therefore the results of study two may be subject to a resultant bias.

Attempts were made to select a sample of subjects from study one to be the subjects in study two at random. However only six of the initial ten subjects identified by random sampling agreed to be interviewed. In order to fill the remaining six positions, a further ten potential subjects were then selected of whom three agreed to be interviewed. The final position was determined by a similar procedure. The final ten subjects were therefore, to a degree, self-selecting and study two is exposed to the risk of selection bias.

7.2.3 Potential presence of interviewer attribute effects in study two

Interviewer attribute effects occur when the personal attributes of the interviewer affect interviewee responses and introduce bias into their responses. In the present research, the decision for subjects to be interviewed by the person known to have designed the study material may have resulted in the presence of such effects. The reasons for this decision are detailed in the section 'Design' in the report of study two (See page 151).

7.2.4 On the internal validity of the logistic regression analysis used to test hypothesis four

The logistic regression analysis used to investigate hypothesis four in studies one, three and four attempts to identify a causal relationship between three factors and subjects' resultant choice to view a feedback type. As discussed in section 'Research methods: Reliability and validity' (See page 96), internal validity to the degree obtained via a true experimental design was not possible. Neither a pre-test nor assigning subjects randomly to a control and an experimental group were possible given that this research took place in an authentic context.

It would have been possible to improve the internal validity of the logistic regression analysis by relaxing this requirement for authenticity. For example, the multiple choice tests could have been reorganised and released to students periodically in order to conduct pre-tests and post-tests of subjects' behaviour surrounding an interventional change such as changing the focus of questions. Artificial subject groups could have been created to represent control and experimental groups as opposed to the use of authentic cohorts.

A trade-off exists between the internal validity arising from increasingly experimental conditionals and the external validity resulting from the authenticity of the learning environment.

7.2.4.1 The presence of potentially confounding factors in the logistic regression analysis used to test hypothesis four

A further consideration relating to the internal validity of the logistic regression analysis is the possible presence of confounding factors. Such factors are identified in the 'Reliability and validity' sections of the reports of studies one, three and four (See pages 129, 186, 211) and four factors are present in each study. The first factor is '#I7 Error type'. The remaining three are affective factors internal to the subject: factors '#I1 Learner expectations', '#I2 Response certitude' and '#I3 Interpretation of task'.

As discussed in the section 'Formation of hypotheses' in the report of study one (See page 126), the inclusion of these four factors into the logistic regression analysis was not possible due to the decision to conduct the study in an authentic educational context. Investigating error type requires the creation of a taxonomy of error types and, subsequently, precise control of the learning environment. The incorporation of factors '#I1 Learner expectations', '#I2 Response certitude' and '#I3 task' into the user interface adopted in the technology based observations would have complicated the learning activities to an unacceptable degree given the decision to conduct this research within an authentic learning environment.

7.2.5 On the role of the semi-structured interviews

During the final analysis, it became apparent that the semi-structured interviews were more valuable to this analysis than had been predicted in the initial planning stage. At the outset of the research, it was predicted that the primary role of these interviews would be to guide the direction of the technology based observations and that it would be the latter that would provide the substantive data. While this is largely the case, the findings from the semi-

structured interviews play an important role and additional semi-structured interviews with subjects participating in studies three and four may have been valuable.

7.3 LIMITATIONS

There are a number of limitations relating to the generalizability of this research and to the coverage of factors potentially affecting learners' interactions with TBI feedback and these are discussed below. However, the primary limitation is that this research does not directly address the question of what format of TBI feedback is most effective.

This is pertinent in that it compromises the immediate applicability of the results to practice. Good educational research should make a valuable contribution to practice (Pring 2004: 1-4) and the results of this research do imply a number of recommendations for good practice. These are provided in the appendices to this thesis in the section "Appendix two: Recommendations for practitioners". However a comprehensive recommendation for the format of TBI feedback, based on evidence of relative effectiveness, is not proposed. Rather, the results of this research have implications for research aiming to answer this larger question.

Other limitations of this research are as follows:

7.3.1 Generalizability

There are a number of limits to the generalizability of this research.

Subject area

The context of this research is postgraduate students at one institution studying three items of TBI across three subject areas. The research will not claim that any findings are generally applicable to other subjects.

Focus on TBI feedback

As discussed in the review of prior research, numerous types of feedback exist in the educational context. The results of this experiment are limited to TBI feedback.

Context

The investigations are performed in the context of technology based instruction and any results are not intended to be generalized to other educational contexts.

Generalizability to other institutions

While results may be indicative for comparable institutions, the generalizability of the results may be compromised by the fact that this research was conducted at a single institution.

7.3.2 Additional possible factors affecting learner preferences

This research did not investigate all factors that potentially affect whether learners choose to view a TBI feedback type. This research identifies 13 such factors. Five were included in the analyses of the data obtained from the technology based observations. 10 of these factors arose in the semi-structured interviews comprising study two. Three factors were not investigated further.

7.4 SUGGESTIONS FOR FUTURE RESEARCH

7.4.1 Further research into learners' preferences for TBI feedback

A comprehensive investigation of learners' preferences is beyond the scope of a single study. Research into the following areas is suggested.

7.4.1.1 Conduct research with different cohorts and different subjects

As Pring (2004: 33) states in relation to observational research, the more observations that are made the greater the validity. Over the past 40 years, and in the 1970s and 1980s in particular, a large number of studies have been conducted that attempt to identify relationships between variations in the format of TBI feedback and performance scores in subsequent tests. These studies have been conducted in a wide range of contexts involving a range of subject matters. This research adopts a novel approach to the issue of the format of TBI feedback and the findings will become more valuable if the results are replicated in a range of contexts.

7.4.1.2 Test findings under more experimental conditions

As discussed in the section 'The final research design' under 'Research design' (See page 80), a deductive experimental design was not adopted in this research due to a lack of prior research in this area and a subsequent inability to form robust hypotheses. However, the inductive exploratory approach taken in this research reveals results and hypotheses that can be subject to investigation under more experimental conditions. Such research would address issues relating to the internal validity of the logistic regression analyses used to investigate hypothesis four in studies one, three and four.

7.4.1.3 Test potential factors identified in this research but that were not investigated

This research identified 13 factors that potentially affect whether learners choose to view a feedback type. Eight of these factors were not investigated in the technology based observations. The effects of factors '#E3 time limit' and '#E4 location' should be priorities for future research. Higher education is witnessing significant growth in the delivery of programmes via blended learning in which a proportion of learning is delivered online and a proportion is delivered face to face (Allen and Seaman 2010). Educators involved in such programmes need to make decisions regarding the instruction that should be delivered online and that which should be delivered face-to-face.

This research suggests that the time and location in which learners study technology based instruction impacts on the frequency at which they elect to view elaborative feedback provided in response to multiple choice questions. If this is the case, educators should consider delivering such questions online rather than in the classroom.

The effect of factor ‘#E1 assessment’ also remains unclear. Subjects in study two of this research stated three different scenarios affecting whether they chose to view the TBI feedback types. First, if the activity constituted an assessed piece of coursework, second if the activity constituted an assessed piece of work and was held in exam type conditions and third, if the activity itself is not assessed but a direct link exists between the activity and a subsequent assessment. Only the latter was investigated further in this research. Investigation of the former two scenarios may produce valuable insights into the dynamic between assessment and the format of TBI feedback that learners prefer.

7.4.2 Research into the effectiveness of different TBI feedback formats that controls for learners’ interactions with TBI feedback and the factors that affect these interactions.

As discussed in the section, ‘Prior research on the taxonomy of TBI feedback types’ in the ‘literature review’ (See page 46), prior research investigating the impact of different formats of TBI feedback on learners’ subsequent performance has been inconclusive and frequently contradictory.

The model proposed in this research contends that factors at the question level, learner level and environmental level affect learners’ interactions with TBI feedback. As discussed in the section “Implications of the results”(See page 265), it can be hypothesized that the effectiveness of a TBI feedback attribute depends on whether it is viewed by a learner and therefore that these factors affect the effectiveness of TBI feedback and that these factors account for the inconclusive and contradictory results of prior research.

Future research into the effectiveness of TBI feedback that controls for these factors may result in a more consistent pattern of results. In addition, a meta-analysis of prior research that investigates the effect of these factors may reveal the extent to which these they account for previous inconclusive and contradictory results.

Together, these two approaches to future research may establish the format of feedback that should be provided to learners within technology based instruction.

8. APPENDICES

APPENDIX ONE: GLOSSARY

The key terminology adopted in this research are defined as follows:

α

' α ' represents the proportion of questions across a multiple question based activity in which a feedback type is selected.

Click

A click is the fundamental unit of data gathered in the quantitative studies in this research. A click is registered when a subject selects one of the target items in the technology based instruction. Such items include 'questions', 'feedback items' and 'software settings'.

E-learning

The use of information and communications technologies (ICT) in learning and teaching.

Educational feedback

"Any communication or procedure given to inform a learner of the accuracy of a response, usually to an instructional question." (Mory 2003 citing Carter, 1984; Cohen, 1985; Kulhavy, 1977; Sales, 1993: 745)

Elaborative feedback (Also known as extended feedback)

This is feedback that provides information beyond whether a response was correct or incorrect.

EX (Explanatory feedback) type feedback

This type of feedback presents an explanation as to why a response is correct or incorrect.

EI (Extra-instructional) type feedback

This type of feedback presents additional information that has not been supplied to the learners previously.

Format

This research proposes a taxonomy of TBI feedback consisting of six feedback types, load and two conditions regarding the timing at which feedback is provided. The format of TBI feedback refers to a particular selection of the attributes in this taxonomy.

IB (Instruction based) type feedback

This type of feedback presents relevant information from the course materials being studied.

Instruction

“Any deliberate arrangement of events to facilitate a learner’s acquisition of some goal.” (Driscoll 2000: 25)

Instructional feedback

Instructional feedback is feedback provided in the course of instruction (Instruction is defined above).

KCR (Knowledge of correct response) type feedback

This type of feedback presents the correct response to a question.

Knowledge gap

The gap between what was understood and what had been required to be understood in order to have responded to a question correctly.

KR (Knowledge of result) type feedback

This type of feedback informs the student as to whether their response was correct.

Learning

“A persisting change in performance or performance potential that results from experience and interaction with the world.” (Driscoll 2000)

NF (No feedback)

The condition in which no feedback is provided to the learner following their response to a question.

Objective question

“Objective tests require a user to choose or provide a response to a question whose correct answer is predetermined.” (CAA 2002)

Question focus

In this research, ‘question focus’ refers to whether a question tests recall or application of the material.

Subject matter

In this research, subject matter refers to whether a question quantitative or qualitative material.

Test

In this research, ‘test’ refers to a series of multiple choice questions.

TBI (Technology based instruction)

Technology based instruction is defined in this research as instruction delivered by computer. Also known as Technology Assisted Instruction (TAI), Computer Assisted Instruction (CAI), Computer Based Instruction (CBI) or Computer Based Training (CBT).

TBI feedback (Technology based instructional feedback)

Technology based instructional feedback is defined as instructional feedback provided in the course of technology based instruction.

Technology based observation

A data gathering technique in which computers track and record subjects' actions.

Technology enhanced learning (TEL)

The use of information and communications technologies (ICT) in learning and teaching. Also known as 'e-learning', 'learning technology' or 'educational technology'.

APPENDIX TWO: RECOMMENDATIONS FOR PRACTITIONERS

At the start of this thesis it was noted that the format of TBI feedback that should be provided to learners was unclear. While this research is not able to provide a comprehensive answer to this question, a number of recommendations to practitioners relating to the content, timing and delivery of TBI feedback are implied.

Recommendation one: Provide TBI feedback immediately following learners response to a question

In this research, learners expressed a clear preference for immediate feedback. 95.22 per cent of subjects chose to view feedback immediately following their response to a question as opposed to viewing feedback at the end of the exercise.

Subjects provided an explanation; they considered immediate feedback to be more efficient as they are already holding the detail of the question in mind at the point feedback is received. Theoretical support for this view is provided by the framework adopted in this research, specifically schema theory. This states that in order for new knowledge to be acquired, relevant schemas need to be activated and loaded into working memory (Derry 1996: 167).

Given this, and the fact that research is yet to establish the relative effectiveness of immediate and delayed feedback (See page 52), it seems difficult to justify the provision of delayed feedback.

Recommendation two: Focus the content of elaborative TBI feedback on anticipated knowledge gaps

This research proposes a model of learners' engagement with TBI feedback in which learners first pursue affirmation as to the correctness of their response by electing to view KR type feedback. Should learners proceed to view the elaborative feedback types, they first elect to view KCR type feedback and then prioritize viewing those feedback types focused most strongly on the knowledge gap identified by a question.

As discussed in the previous chapter 'Discussion' (See page 250), this is explained by a focus on the pursuit of closure at the question level. Cognitive load theory offers some explanation for this focus by pointing to the limitations of working memory (See page 253).

Should educators wish to encourage connections to be made between a question and the higher activity or course level objectives then additional instruction should be provided for this purpose at a later stage.

Recommendation three: Allow learners to revisit questions

In the studies comprising this research between 37.69 per cent and 86.30 per cent of subjects revisited questions. When designing tests in TBI, educators make a choice whether to allow learners to revisit question. For example,

the online learning systems used to deliver multiple choice tests within higher education commonly allow educators to make this choice.

In this research, subjects' demonstrated a strong focus on achieving closure or, equivalently, resolving perturbations at the question level and it is assumed here that when subjects' elect to revisit a previously viewed question, they do so as part of this process. Educators can assist in this process by allowing learners to revisit questions.

Recommendation four: Provide elaborative TBI feedback to correct responses

Elaborative feedback is not always provided to correct responses to multiple choice questions within TBI. This research found that, while responding to a question incorrectly dramatically increased the frequency at which subjects viewed the elaborative feedback types, subjects did view elaborative feedback following a correct response. Study two provided an explanation for this, subjects stated they had guessed the correct response on the occasions in which they did so. Therefore, this research recommends that educators allow for the presence of guessing and provide meaningful feedback to correct responses.

Recommendation five: Be conscious of TBI feedback load

Feedback load, that is the volume of feedback presented, emerged as a clear issue in this research. Seven of the ten subjects in study two made explicit reference to load during their interviews. These subjects held negative views towards what they considered to be a high feedback load and gave three reasons for this. First, subjects sought a time efficient route towards closure. Second, subjects stated that a high feedback load dampens a sense of progress through the activity. Third, subjects found higher volumes of feedback 'overwhelming'. These results are supported by Pridemore and Klein (1995) who found that subjects who received KCR plus EX type feedback were less likely to have enjoyed the lesson and less likely to feel they had 'learned a lot' than those receiving KCR type feedback alone.

This poses a dilemma to educators. The occasions on which learners elect to view high feedback loads likely result in facilitating the deep learning orientated behaviour that is the aim of most educators. However, there is a tension between providing a feedback load sufficient to promote such experiences and satisfying learners' general preference for low feedback load.

This research recommends that educators be mindful of this dilemma when designing TBI feedback. One possible solution is to make elaborative TBI feedback beyond KCR available but optional.

Recommendation six: Consider the impact of 'time' and 'place' when you consider elaborative TBI feedback to be important

Study one observed subjects studying the 'Ethics and Strategy' exercise' item of TBI together in a computer room with a time limit. Despite being given significantly more time than required, those subjects from study one interviewed in study two made frequent reference to both the presence of a time limit and the location of the

activity being a factor affecting whether they choose to view the feedback types. These subjects stated they would be more likely to view the elaborative feedback types when taking the test at a time and place of their choosing. Studies three and four provided limited, circumstantial support for this. Subjects studying the “Accounting Primer” and “Pre-study Mathematics” courses did so at a time and place of their choosing and without a time limit. The frequency at which subjects elected to view the elaborative feedback types for these two items of TBI was noticeably higher than for the ‘Ethics and Strategy’ exercise.

This research does not provide sufficient evidence to draw firm conclusions. However, it is recommended here that educators be mindful of the impact of time and place when conducting tests in which the provision of elaborative feedback is considered important to achieving the goals of the activity.

Recommendation seven: Provide elaborative feedback beyond KCR type feedback for questions that test application of the study material

Generally, elaborative feedback is more complex than verificational feedback and its development requires educators to make a significantly higher investment of time and resources. Such feedback also requires a significantly higher investment of time and effort by the learner. An important question for educators is whether the resulting learning benefits justify this investment.

As stated in a previous section Kulhavy et al. (1985: 285) explicitly address the issue of feedback efficiency using the following measure of feedback efficiency.

$$\text{Feedback efficiency} = \frac{\text{correct retention percentage score}}{\text{square root of feedback study time}}$$

Kulhavy et al. conclude that elaborative feedback beyond KCR type feedback is not efficient and does not represent a worthwhile investment of learners’ time. Dempsey and Litchfield (1993), adopting the same measure of efficiency, also attempt to determine the relative efficiencies of different feedback types. As per Kulhavy et al. this study established that students spent more time studying the elaborate types of feedback but also found no difference in performance in the retention test across the four conditions. Using the measure of feedback efficiency above, they therefore concluded that KCR feedback is the most efficient type of feedback among the four.

This research does not provide evidence to support or counter the concept of feedback efficiency but cautions against adopting the findings of these two studies too readily as the level of feedback efficiency may depend on the nature of the TBI material. This research found that subjects viewed elaborative feedback types beyond KCR significantly more frequently when questions tested application as opposed to recall of the subject matter. In addition, subjects in study two suggested that they would be more likely to view these feedback types if the subject matter was quantitative rather than qualitative.

APPENDIX THREE:

TOPIC GUIDE FOR SEMI STRUCTURED INTERVIEWS

Introductory statement

My name is David Lefevre. I am conducting this research as part of my doctoral studies at the Imperial College Business School. The aim of this research is to identify students' preferences for feedback in the type of activity you have just taken. This will help us improve the design of such activities and increase our understanding of the ways in which students such as yourself prefer to study.

You were selected for this interview at random and for no other reason. All information you provide during this interview will be confidential. Your participation in his research will be anonymous. Identifying factors, such as your name, will not be recorded and as such will not appear in any documentation that results from this research. Participation in this interview is voluntary, you may leave at any time.

Do you have any questions before we begin?

(A screenshot of the user interface is given to the respondent for reference)

Questions

1. How often did you select the feedback item "Did I answer correctly"?

- For all the questions
- For the majority of the questions
- For approximately half the questions
- For less than half of the questions
- For none of the questions

2. What was the primary reason you chose to select this item?

3. Where there any other reasons you chose to select this item?

4. What was the primary reason you chose not to select this item?

5. Where there any other reasons you chose not to select this item?

6. Can you think of any factors that may determine whether you choose, or not choose, to view this feedback item?

Repeat questions 1-6 for the following feedback items:

What was the correct answer?

Book reference

Further explanation

Example

Explanation of terms

7. Would you like to see any additional types of feedback in activities such as this?

8. Do you have any further comments relating to the feedback provided in this activity?

APPENDIX FOUR:

COMPARISON OF PRIOR RESEARCH ON FEEDBACK TAXONOMY

Study	Feedback type / Attribute										Difference?	Summary
	Content						Load	Timing				
	Verificational		Elaborative									
	NF	KR	KCR	EX	IB	EI						
Gilman (1969)	✓	✓	✓	✓							Y	Subjects receiving KCR and EX feedback performed significantly better in a subsequent performance test than those receiving KR or NF. There was no significant difference in performance between learners who received KCR and those who received EI. No significant difference between NF and KR
Hanna (1976)	✓	✓	✓								Y	Provision of either feedback type significantly improved performance compared to the provision of no feedback. No significant difference was found between feedback types.
Roper (1977)	✓	✓	✓								Y	Performance in a post test following the instruction indicated that KCR type feedback is more effective than KR which in turn is more effective the provision of no feedback.
Peeck (1979)			✓		✓						N	No significant difference found.
Phye (1979)							✓				Y	Information beyond that which is necessary has a detrimental effect on post-test performance.
Kulhavy et. al (1985)			✓	✓	✓		✓				N	Providing KCR type feedback alone resulted in higher increase in subsequent performance than providing KCR plus EX type feedback.
Birenbaum and Tatsuoka (1987)		✓	✓		✓						Y	The effectiveness of the three feedback types in correcting non serious learner errors differed significantly The more informative the feedback, the greater the effect.
Andre and Thieman (1988)	✓		✓		✓						Y	Subjects who received KCR or IB type feedback performed significantly better in a post performance test than subjects who did not receive feedback but only for questions testing factual knowledge. No significant difference between KCR and IB.
Clarina, Ross et al. (1991)	✓		✓					✓	✓		Y	Subjects who received any of the feedback types performed better in a post performance test than subjects who did not receive feedback.

Dempsey and Litchfield (1993)		✓				N	No significant difference across feedback o types in scores in subsequent performance test.
Morrison, Ross et al. (1995)	✓		✓		✓	✓	Y KCR type feedback was found to be more e effective than no feedback. Delayed s feedback was more effective than immediate feedback but only for “lower level” learning.
Pridemore and Klein (1995)	✓		✓	✓			Y Subjects did not receive feedback or received e EX type feedback performed significantly s better in a post performance test than subjects who received KCR type feedback.
Dihoff, Brosvic et al. (2003)	✓				✓	✓	Y Subjects receiving immediate feedback e performed significantly better in a post s performance test than subjects who did not receive feedback or received delayed feedback.
Moreno (2004)			✓	✓			Y Significant differences were found in the e effects of the two feedback types on s retention, transfer, perceived difficulty and perceived helpfulness of the software. But not on motivation and interest. Groups presented with IB type feedback scored significantly better in a subsequent performance test.
Pashler, Cepeda et al. (2005)	✓	✓	✓				Y Performance is a post test following the e instruction indicated than KCR is more s effective than KR type feedback which in turn is more effective not providing feedback.

BIBLIOGRAPHY

- Allen, I. E. and J. Seaman (2010). *Class Differences: Online Education in the United States, 2010*, The Sloan Consortium.
- Alliance, A. (2012). "The Twelve Principles of Agile Software ". Retrieved 29th June 2012, from <http://www.agilealliance.org/the-alliance/the-agile-manifesto/the-twelve-principles-of-agile-software/>.
- Anderson, R. C. and P. D. Pearson (1984). A schema-theoretic view of basic processes in reading comprehension. *Handbook of reading research*. P. D. Pearson. New Jersey, Lawrence Erlbaum Associates: 255-291.
- Anderson, R. C., R. J. Spiro, et al. (1978). "Schemata as scaffolding for the representation of information in connected discourse." *American Educational Research Journal* **15**(3): 433-440.
- Andre, T. and A. Thieman (1988). "Level of adjunct question, type of feedback, and learning concepts by reading." *Contemporary Educational Psychology* **13**(3): 296-307.
- Askew, S. and C. Lodge (2000). Gifts, ping-pong and loops - linking feedback and learning. *Feedback for learning*. S. Askew. London, Routledge.
- AUT (2005). Packing them in - The student-to-staff ratio in UK higher education, Association Of University Teachers.
- Bangert-Drowns, R. L., C.-L. C. Kulik, et al. (1991). "The Instructional Effect of Feedback in Test-Like Events." *Review of Educational Research* **61**(2): 213-238.
- Birenbaum, M. and K. K. Tatsuoka (1987). "Effects of "On-Line" Test Feedback on the Seriousness of Subsequent Errors." *Journal of Educational Measurement* **24**(2): 145-155.
- Bloom, B. S. (1984). "The 2-sigma problem." *Educational researcher* **13**(6): 4-16.
- Bloom, B. S., M. D. Engelhart, et al. (1956). *Taxonomy of Education Objectives the Classification of Educational Goals: Handbook I Cognitive Domain*. New York, David McKay Company, Inc.
- Boehler, M., D. Rogers, et al. (2006). "An investigation of medical student reactions to feedback: a randomised controlled trial." *Medical education* **40**(8): 746-749.
- Brackbill, Y. and M. S. Kappy (1962). "Delay of reinforcement and retention." *The Journal of comparative and physiological psychology* **55**: 14-18.
- Bradbard, D., D. Parker, et al. (2004). "An Alternate Multiple-Choice Scoring Procedure in a Macroeconomics Course." *Decision sciences journal of innovative education* **2**(1): 11.
- Brown, G., J. Bull, et al. (1997). *Assessing student learning in higher education*. London, UK, Routledge.
- Browne, T., R. Hewitt, et al. (2010). 2010 Survey of Technology Enhanced Learning for higher education in the UK, UCISA.

- Bruner, J. S., J. J. Goodnow, et al. (1967). Study of Thinking. New York.
- Bryman, A. and E. Bell (2003). Business research methods. Oxford, OUP
- Buckles, S. and J. Siegfried (2006). "Using Multiple-Choice Questions to Evaluate In-Depth Learning of Economics." The Journal of economic education **37**(1): 48.
- Butler, A. C., J. D. Karpicke, et al. (2007). "The effect of type and timing of feedback on learning from multiple-choice tests." Journal of Experimental Psychology: Applied **13**(4): 273-281.
- Butler, D. L. and P. H. Winne (1995). "Feedback and Self-Regulated Learning: A Theoretical Synthesis." Review of Educational Research **65**(3): 245-281
- CAA (2002). "Objective Test Design." Retrieved 12/12/08, from http://www.caacentre.ac.uk/resources/objective_tests/index.shtml.
- Chandler, P. and J. Sweller (1991). "Cognitive Load Theory and the Format of Instruction." Cognition and Instruction **8**(4): 293-332.
- Clarina, R. B., S. M. Ross, et al. (1991). "The effects of different feedback strategies using computer-administered multiple-choice questions as instruction." Educational Technology Research and Development **39**(2): 5-17.
- Clark, R. E. (1983). "Reconsidering Research on Learning from Media." Review of Educational Research **53**: 445-459.
- Clark, R. E., K. Yates, et al. (2010). An Analysis of the Failure of Electronic Media and Discovery-Based Learning. Handbook of Improving Performance in the Workplace: Instructional Design and Training Delivery, John Wiley & Sons, Inc.: 263-297.
- Coffield, F., D. Moseley, et al. (2004). Learning styles and pedagogy in post-16 learning A systematic and critical review, Learning and Skills Research Centre.
- Cohen, L., L. Manion, et al. (2007). Research Methods In Education. Abingdon, Routledge.
- CST (2006). Universal ethical code for scientists. C. f. S. a. Technology, London.
- Dempsey, J. V., M. P. Driscoll, et al. (1993). Text-based feedback. Interactive instruction and feedback. J. V. Dempsey and G. C. Sales. Englewood Cliffs, NJ, Educational Technology Publications: 21–54.
- Dempsey, J. V. and B. C. Litchfield (1993). "Feedback, retention, discrimination error, and feedback study time." Journal of research on computing in education **25**(3): 303.
- Dempsey, J. V. and S. U. Wager (1988). "A Taxonomy for the Timing of Feedback in Computer-Based Instruction." Educational Technology **28**(10): 20-25
- Derry, S. J. (1996). "Cognitive schema theory in the constructivist debate." Educational Psychologist **31**(3-4): 163-174.
- Dihoff, R. E., G. M. Brosvic, et al. (2003). "The Role of Feedback During Academic Testing: The Delay Retention Effect Revisited." The Psychological record **53**(4): 533.
- Driscoll, M. P. (2000). Psychology of Learning for Instruction. Needham Heights, MA, Allyn & Bacon.
- Driscoll, M. P. and M. Tessmer (1985). Applications of the concept tree and rational set generator for coordinate concept learning. The annual meeting of the American Educational Research Association. Chicago, IL. .

- Duffy, T. M. and D. J. Cunningham (1996). Constructivism: Implications for the Design and Delivery of Instruction. Handbook of Research for Educational Communications and Technology. D. H. Jonassen. New York, Simon & Shuster Macmillan: 170- 198.
- Farwell, R. (2002). Higher Education Provision and the Change to a Mass System. Access, Participation and Higher Education: Policy and Practice. A. Hayton and A. Paczuska: 151.
- Gagne, R. M. (1985). The conditions of learning. New York.
- Gibbs, G. R. (2002). Qualitative data analysis: Explorations with NVivo. Maidenhead, England., Open University Press.
- Gilman, D. A. (1969). " Comparison of several feedback methods for correcting errors by computer-assisted instruction." Journal of Educational Psychology **6**(6): 503-508.
- Glaser, B. G. and A. L. Strauss (1967). The Discovery of Grounded Theory. Chicago, IL, Aldane.
- Goodman, J. S. (1998). "The Interactive Effects of Task and External Feedback on Practice Performance and Learning." Organizational Behavior and human decision processes **76**(3): 223–252.
- Grant, H. and C. S. Dweck (2003). " Clarifying Achievement Goals and Their Impact." Journal of Personality and Social Psychology **85**(3).
- Greenaway, D. and M. Haynes (2003). "Funding Higher Education in the UK: The Role of Fees and Loans." The economic journal **113**(485): 150-166.
- Guest, G., A. Bunce, et al. (2006). "How Many Interviews Are Enough? An Experiment with Data Saturation and Variability." Field Methods **18**(1): 59-82.
- Handley, K. (2002). Comparison of novice and expert learner's perception of instructional feedback in computer-based training to develop managerial problem-solving skills. HCT2002 Workshop – Tools for thought: Communication and Learning Through Digital Technology, University of Sussex, Brighton, UK. .
- Handley, K. (2003). Achieving Closure: The Function of Feedback in e-Learning Activities. Business School. London, Imperial College. PhD.
- Handley, K. and B. Cox (2007). "Beyond model answers: learners' perceptions of self-assessment materials in e-learning applications." ALT-J. Association for Learning Technology journal **15**(1): 21.
- Handley, K., M. Price, et al. (2011). "Beyond 'doing time': investigating the concept of student engagement with feedback." Oxford Review of Education **37**(4): 543–560.
- Hanna, G. S. (1976). "Effects of Total and Partial Feedback in Multiple-Choice Testing upon Learning." The Journal of Educational Research **69**(5): 202-205.
- Hannafin, M. J. (1992). "Emerging technologies, ISD, and learning environments: Critical perspectives." Educational technology research and development **40**(1): 49.
- Hattie, J. and H. Timperley (2007). "The Power of Feedback." Review of Educational Research **77**,(1): 81–112.
- HEFCE (2010). "National Student Survey." Retrieved 17th August 2011, from <http://www.hefce.ac.uk/news/hefce/2010/nssresult.htm>.
- HESA (2011). Statistics - Students and qualifiers at UK HE institutions.
- Honey, P. and A. Mumford (1992). The Manual of Learning Styles. Maidenhead, Peter Honey Publications.

- Huxham, M. (2007). "Fast and effective feedback: are model answers the answer?" Assessment & Evaluation in Higher Education **32**(6): 601 - 611.
- Ivancevich, J. M. and J. T. McMahon (1982). "The effects of goal setting, external feedback, and self-generated feedback on outcome variables: A field experiment." The Academy of Management journal **25**(2): 359.
- JISC (2008). Tangible Benefits of e-Learning: Does investment yield interest?
- JISC-Infonet (2009). "Tools in a Virtual learning Environment." Retrieved 2/1/2009, from <http://www.jiscinfonet.ac.uk/InfoKits/effective-use-of-VLEs/intro-to-VLEs/introtovle-intro/introtovle-tools>.
- Keller, J. M. (1987). "Development and use of the ARCS model of instructional design." Journal of instructional development **10**(3): 2-10.
- King, P. E., P. Schrodtt, et al. (2009). "The Instructional Feedback Orientation Scale: Conceptualizing and Validating a New Measure for Assessing Perceptions of Instructional Feedback." Communication Education **58**(2): 235-261.
- Kluger, A. N. and A. DeNisi (1996). "The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory." Psychological bulletin **119**(2): 254-284.
- Kolb, D. A. (1984). Experiential learning: experience as the source of learning and development. Englewood Cliffs, NJ, Prentice Hall.
- Kolb, D. A. and R. Fry (1975). Toward an applied theory of experiential learning. Theories of Group Process. C. Cooper. London, John Wiley.
- Kozma, R. B. (1991). "Learning with Media." Review of Educational Research **61**: 179-211.
- Kraiger, K. and E. Jerden (2007). A meta-analytic investigation of learner control: Old findings and new directions. Toward a science of distributed learning. S. M. Fiore and E. Salas. Washington, DC, US, American Psychological Association: 65-90.
- Kulhavy, R. W. (1977). "Feedback in Written Instruction." Review of Educational Research **47**(2): 211-232.
- Kulhavy, R. W. and W. Stock (1989). "Feedback in written instruction: The place of response certitude." Educational psychology review **1**(4): 279.
- Kulhavy, R. W., M. T. White, et al. (1985). "Feedback complexity and corrective efficiency." Contemporary Educational Psychology **10**(3): 285.
- Kulik, J. A. and C.-L. C. Kulik (1988). "Timing of Feedback and Verbal Learning." Review of Educational Research **58**(1): 79-97.
- Laurillard, D. (2002). Rethinking University Teaching: A Conversational Framework for the Effective Use of Learning Technologies.
- Lave, J. and E. Wenger (1991). Situated Learning: Legitimate Peripheral Participation. Cambridge, Cambridge University Press.
- Maxwell, J. A. (1992). "Understanding and Validity in Qualitative Research." Harvard Educational Review **62**(3).
- Mayer, R. E. and R. Moreno (2002). "Aids to computer-based multimedia learning." Learning and instruction **12**(1): 107-119.

- Mayer, R. E., R. Moreno, et al. (1999). "Maximizing constructivist learning from multimedia communications by minimizing cognitive load." Journal of Educational Psychology **91**(4): 638-643.
- Moreno, R. (2004). "Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia." Instructional science **32**(1): 99.
- Morrison, G. R., S. M. Ross, et al. (1995). "The Effects of Feedback and Incentives on Achievement in Computer-Based Instruction." Contemporary Educational Psychology **20**(1): 32.
- Morse, J. M. (2000). "Determining Sample Size." Qualitative Health Research **10**(1): 3-5.
- Mory, E. H. (2003). Feedback Research Revisited. Handbook of research for educational communications and technology, 2nd Edition. D. H. Jonassen. Mahwah, NJ, Lawrence Erlbaum.: 745-783.
- Niemiec, R. P., C. Sikorski, et al. (1996). "Learner-Control Effects: A Review of Reviews and a Meta-Analysis." Journal of educational computing research **15**(2): 157-174.
- Pargetter, R., C. McInnis, et al. (1998). Transition from Secondary to Tertiary: A Performance Study.
- Pashler, H., N. J. Cepeda, et al. (2005). "When does feedback facilitate learning of words." Journal of experimental psychology. Learning, memory, and cognition **31**(1): 3.
- Peeck, J. (1979). "Effects of Differential Feedback on the Answering of Two Types of Questions by Fifth- and Sixth-Graders." British Journal of Educational Psychology **49**: 87-92.
- Phye, G. D. (1979). "The processing of informative feedback about multiple-choice test performance." Contemporary Educational Psychology(4): 381-394.
- Piaget, J. (1967). Six Psychological Studies. New York, John Wiley.
- Pridemore, D. R. and J. D. Klein (1995). "Control of Practice and Level of Feedback in Computer-Based Instruction." Contemporary Educational Psychology **20**(4): 444.
- Pring, R. (2004). Philosophy of Educational Research. London, Continuum.
- Respondus (2011). "About Respondus ". Retrieved 05/07/2011, from <http://www.respondus.com/about/index.shtml>.
- Reynolds, M. (1997). "Learning Styles: A Critique." Management Learning **28**(2): 115-133.
- Ritchie, J. and J. Lewis (2003). Qualitative research practice: A guide for social science students and researchers. London, Sage.
- Robbins, L. C. (1963). The Robbins Report, London: Her Majesty's Stationery Office, Report of the Committee appointed by the Prime Minister under the Chairmanship of Lord Robbins.
- Roper, W. J. (1977). "Feedback in computer assisted instruction." Programmed Learning and Educational Technology **14**(1): 43-49.
- Ross, S. M., G. R. Morrison, et al. (1989). "Uses and effects of learner control of context and instructional support in computer-based instruction." Educational technology research and development **37**(4): 29.
- Sadler, D. R. (1989). "Formative assessment and the design of instructional systems." Instructional science **18**(2): 119-144.

- Sales, G. C. (1993). Adapted and adaptive feedback in technology-based instruction. Interactive instruction and feedback J. V. Dempsey and G. C. Sales. Englewood Cliffs, NJ, Educational Technology: 159–175.
- Schimmel, B. J. (1988). Providing meaningful feedback in courseware. Instructional designs for microcomputer courseware. D. H. Jonassen, Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc: 183-195.
- Shadish, W. R., T. D. Cook, et al. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston, Houghton Mifflin Company.
- Shute, V. J. (2008). "Focus on Formative Feedback." Review of Educational Research **78**(1): 153-189.
- Simkin, M. and W. Kuechler (2005). "Multiple-Choice Tests and Student Understanding: What Is the Connection?" Decision sciences journal of innovative education **3**(1): 73.
- Skinner, B. F. (1938). The behavior of organisms. New York, Appleton-Century-Crofts.
- Skinner, B. F. (1968). The Technology of Teaching. New York, Appleton-Century-Crofts.
- Steneck, N., M. S. Anderson, et al. (2012). Research Integrity: Social and Behavioural Sciences. Research Integrity. C. Travers. London, Epigeum.
- Steneck, N., T. Mayer, et al. (2010). Singapore Statement on Research Integrity. 2nd World Conference on Research Integrity.
- Sweller, J. (1988). "Cognitive load during problem solving: Effects on learning." Cognitive science **12**(2): 257-285.
- Sweller, J., J. van Merriënboer, et al. (1998). "Cognitive Architecture and Instructional Design." Educational psychology review **10**(3): 251-296.
- Tait, H., N. J. Entwistle, et al. (1998). ASSIST: A re-conceptualisation of the approaches to studying inventory. Improving student learning: Improving students as learners. C. Rust. Oxford, Oxford Centre for Staff and Learning Development, Oxford Brookes University: 262–271.
- Tatsuoka, K. K. (1984). "Changes in error types over learning stages." Journal of Educational Psychology **76**(1): 120-129.
- Trowbridge, M. H. and H. Cason (1932). "An experimental study of Thorndike's theory of learning." Journal of General Psychology **7**: 245-260.
- Udacity (2012). "Introduction to Computer Science (cs101)." Retrieved 31.12.2012, from <http://www.udacity.com/overview/Course/cs101/CourseRev/apr2012>.
- Valcke, M. (2002). "Cognitive load: updating the theory?" Learning and instruction **12**(1): 147-154.
- Von Glasersfeld, E. (1989). "Cognition, construction of knowledge, and teaching." Synthese **80**(1): 121.
- Vygotsky, L. S. (1978). Mind in Society: Development of Higher Psychological Processes. Cambridge, MA, Harvard University Press.
- Wager, W. and S. Wager (1985). "Presenting Questions, Processing Responses, and Providing Feedback in CAI." Journal of Instructional Development **8**(4): 2-8.
- Walker, M. (2009). "An investigation into written comments on assignments: do students find them usable?" Assessment & Evaluation in Higher Education **34**(1): 67–78.

- Weaver, M. (2006). "Do students value feedback? Student perceptions of tutors' written responses." Assessment & Evaluation in Higher Education **31**: 379-394.
- Wiley (2009). "WileyPlus." Retrieved 1/1/2009, from <http://he-cda.wiley.com/WileyCDA/Section.rdr?id=101003>.
- Williams, J. B. (2006). "Assertion-reason multiple-choice testing as a tool for deep learning: a qualitative analysis." Assessment & Evaluation in Higher Education **31**(3): 287 - 301.
- Yin, R. (2002). Case Study Research: Design and Methods, Third Edition, Applied Social Research Methods Series, Vol 5, Sage Publications, Inc.