



## **University of Bradford eThesis**

This thesis is hosted in [Bradford Scholars](#) – The University of Bradford Open Access repository. Visit the repository for full metadata or to contact the repository team



© University of Bradford. This work is licenced for reuse under a [Creative Commons Licence](#).

# INVESTIGATING ASSESSMENT IN HIGHER EDUCATION: DEMYSTIFYING MARKING AND GRADING TO REVEAL EXPERT PRACTICES

A phenomenological analysis of marking and grading practices of  
novice and experienced health academics

**Fiona Samantha MEDDINGS**

submitted for the degree of  
Doctor of Philosophy

School of Allied Health Professions & Midwifery  
Faculty of Health Studies

University of Bradford

2017

# ABSTRACT

Fiona Samantha MEDDINGS

## **Investigating Assessment in Higher Education: Demystifying Marking and Grading to Reveal Expert Practices**

A phenomenological analysis of marking and grading practices of novice and experienced health academics

**Key words:** Marking, grading, assessment, lecturers, higher education, protocol analysis, Cognitive Interviewing, health studies

This thesis is focused on exploring marking and grading in higher education. Using a phenomenological approach 26 interviews were undertaken with a heterogeneous sample of health academics at four universities. The sample included novice lecturers with two years or less experience in the academy to those with academic careers spanning more than 20 years. Two interview methods were utilised, Protocol Analysis (PA) and Cognitive Interviewing (CI). Protocol Analysis enables close contact with the marking and grading, in the moment, whilst Cognitive Interviewing is a novel method for exploring lecturer practices in higher education. Analysis was completed by applying a modified framework analysis to both data sets, facilitating synthesis of the two series of research findings. A wealth of rich data was gathered which resulted in close exploration of marking and grading practices, with the production of corroboratory evidence for issues previously identified on these phenomena. Close connection as an insider practitioner researcher has enabled close exploration and the gaining of new insights into practice, resulting in the identification of previously unexplored areas. My original contributions to knowledge in this area are: identifying the messiness of marking and grading and troublesome knowledge, the 'rubric paradox', importance of communities of practice, dual identity of health academics, working environments, experience recast as expertise, and using current interview techniques (PA and CI) for supporting continuing professional

development. This thesis develops these themes suggesting ways in which they could impact upon contemporary marking and grading practice.

# ACKNOWLEDGEMENTS

*“There are no limitations, just dreams to be fulfilled”*

Joel Brown

I would like to thank many many people who have been there to support me on the very long journey toward completion and submission of this thesis.

My husband Kevin, for his never ending optimism and enduring faith, that this project, was something that I could and indeed I should do. He found any way possible to facilitate me getting to completion, including removing the need to procrastinate over what we might have for tea. This journey turned him into a great cook!

My daughter Eleanor – who has been on her own academic journey, I hope that I have stood as an example of what can be done if only you set your gaze high enough.

My Supervisor, Dr Sean Walton who has been there right from the outset, you too have shown an enduring faith which when I doubted myself, was there to guide me back to the path. A special mention of Mr Peter Hughes, my primary supervisor, who joined my supervision team not long before I completed my transfer from MPhil to PhD, he has been an exceptional support, your keen eye for detail, your astute questioning style have all been valued. However, over and above that was your faith that this was a project worth doing and that I could do it. I had a 3<sup>rd</sup> member of my supervision team, Professor Maryann Hardy who has always been there as a support should she be needed.

My colleagues in the School of Midwifery & Reproductive Health deserve a special mention. Without them, this project and thesis would never have had the opportunity to flourish, as even when workload was unbearable, they stood by me trying to protect time and space for me to study.

I would also like to mention the PhD student journey which was set on course by Professor Hardy who had a vision for what post graduate students needed, and set about to build it. The result was a successful and

supportive community of students, in a small enclave which was the Faculty of Health Studies PhD student room. Things change, but what remained was the support of 3 colleagues with whom I became a member of a learning set. In its infancy the learning set provided an opportunity for us to learn to talk about our ideas, to articulate them using research speak and to question one another when we were unsure. In the latter stages as we all went separate ways to complete data collection and then writing, our meetings were less frequent, but nonetheless remained a great source of convivial challenge. My thanks therefore go out to my set members Dr Rose Peacock (PhD), Dr Colin Ayre (PhD) and Dr Shehla Khalid (PhD). I also latterly need to thank another research student and colleague Laura Iraine Green who has inspired me to start and keep writing, our occasional writing retreats have been just what I needed.

Finally thanks must go to the interviewees, peers and colleagues at my own and external institutions who have brought to life the inner world of marking and grading which would otherwise have remained hidden. The rich source of data collected has served to highlight and illustrate the complexities of a taken for granted academic activity, which has far reaching consequences for students, departments, faculties, institutions and society as a whole. Findings, as shaped here will perhaps become a catalyst for change.

# TABLE OF CONTENTS

<b>ABSTRACT</b> .....	<b>I</b>
<b>ACKNOWLEDGEMENTS</b> .....	<b>III</b>
<b>TABLE OF CONTENTS</b> .....	<b>V</b>
<b>TABLE OF TABLES</b> .....	<b>XIV</b>
<b>TABLE OF FIGURES</b> .....	<b>XV</b>
<b>GLOSSARY OF TERMS</b> .....	<b>XVI</b>
<b>PROLOGUE</b> .....	<b>XVII</b>
<b>INTRODUCTION</b> .....	<b>XIX</b>
<i>THESIS STRUCTURE</i> .....	<i>xix</i>
<i>SETTING THE SCENE</i> .....	<i>xxi</i>
<b>CHAPTER 1 LITERATURE REVIEW</b> .....	<b>1</b>
<b>SECTION 1</b> .....	<b>3</b>
1.1.1 MARKING OR GRADING CRITERIA, AND RUBRICS .....	3
1.1.2 NORMATIVE AND OR CRITERION REFERENCING .....	7
1.1.3 COMPLETING MARKING AND GRADING AS AN ACTIVITY .....	10
1.1.4 ASSESSMENT LITERACY OF STUDENTS AND STAFF .....	10
<b>SECTION 2</b> .....	<b>14</b>
1.2.1. RELEVANCE OF COMMUNITIES OF PRACTICE FOR ACHIEVING MARKING AND GRADING .....	14
1.2.2. TACIT BELIEFS IN RELATION TO MARKING AND GRADING .....	16
1.2.3. EXPERTNESS AND EXPERT PRACTICE .....	19
<i>Summary</i> .....	27
<b>CHAPTER 2 PHILOSOPHY AND METHODOLOGY</b> .....	<b>29</b>
<i>PHILOSOPHICAL ORIENTATION</i> .....	29
<b>2.1 ONTOLOGY AND EPISTEMOLOGY</b> .....	<b>29</b>
<b>2.2 PRACTICE OF MARKING AND GRADING</b> .....	<b>31</b>
<b>2.3 AIMS, OBJECTIVES AND QUESTIONS</b> .....	<b>32</b>
2.3.1 AIMS .....	33

2.3.2	RESEARCH OBJECTIVES .....	33
2.3.3	RESEARCH QUESTIONS .....	33
	<i>METHODOLOGICAL STANCE</i> .....	33
<b>2.4</b>	<b>RESEARCH APPROACH .....</b>	<b>34</b>
<b>2.5</b>	<b>METHODOLOGICAL MEANS .....</b>	<b>34</b>
2.5.1	PHENOMENOLOGY .....	35
2.5.2	HERMENEUTIC PHENOMENOLOGY .....	36
2.5.3	ACTIVE INTERVIEWING .....	37
2.5.4	RESEARCHER REFLEXIVITY .....	39
2.5.5	METHODOLOGY VS. METHOD.....	40
2.5.6	INTERVIEW SAMPLE .....	41
2.5.7	PRACTITIONER OR INSIDER RESEARCH.....	42
2.5.8	RESEARCH SITE AND POPULATION .....	43
<b>2.6</b>	<b>ETHICAL CONSIDERATIONS .....</b>	<b>43</b>
2.6.1	RESEARCH ETHICS .....	43
2.6.2	LECTURERS/INTERVIEWEES.....	44
2.6.3	GAINING CONSENT .....	44
2.6.4	POPULATION, RECRUITMENT, SAMPLING .....	45
2.6.5	DATA COLLECTION .....	48
2.6.6	CONFIDENTIALITY .....	50
2.6.7	ETHICS OF PRACTITIONER RESEARCH .....	52
	<i>APPROACH FOR DATA ANALYSIS</i> .....	53
<b>2.7</b>	<b>INITIAL THOUGHTS ON FINDINGS .....</b>	<b>53</b>
<b>2.8</b>	<b>APPROACH FOR ANALYSING INTERVIEW DATA .....</b>	<b>53</b>
2.8.1	EMPLOYING FRAMEWORK ANALYSIS .....	55
2.8.2	CI CODING AND THEME DEVELOPMENT.....	56
<b>2.9</b>	<b>IDENTIFYING AN ANALYTICAL APPROACH FOR PA INTERVIEWS.....</b>	<b>61</b>
<b>2.10</b>	<b>PA CODING AND THEME DEVELOPMENT .....</b>	<b>62</b>
<b>2.11</b>	<b>DATA COLLECTION LIMITATIONS .....</b>	<b>63</b>



<b>2.12</b>	<b>SUMMARY</b> .....	<b>64</b>
<b>CHAPTER 3</b>	<b>COGNITIVE INTERVIEWING METHOD</b> .....	<b>66</b>
<b>3.1</b>	<b>WHAT IS COGNITIVE INTERVIEWING?</b> .....	<b>66</b>
<b>3.2</b>	<b>PROCEDURE</b> .....	<b>68</b>
3.2.1	ADVANTAGES .....	70
3.2.2	DISADVANTAGES .....	73
3.2.3	COGNITIVE INTERVIEW PROTOCOL.....	74
3.2.4	INTERVIEWER TRAINING .....	75
3.2.5	CONDUCTING COGNITIVE INTERVIEWING INTERVIEWS .....	75
3.2.6	SELECTION OF INTERVIEWEES TO CONTRIBUTE .....	76
<b>3.3</b>	<b>SUMMARY</b> .....	<b>77</b>
<b>CHAPTER 4</b>	<b>COGNITIVE INTERVIEW FINDINGS</b> .....	<b>78</b>
<b>4.1</b>	<b>RESEARCH AIM</b> .....	<b>78</b>
	RESEARCH OBJECTIVES: .....	78
	RESEARCH QUESTIONS: .....	79
<b>4.2</b>	<b>INTERVIEWEE KEY 1:</b> .....	<b>79</b>
<b>4.3</b>	<b>EXPLORATION OF FINDINGS</b> .....	<b>80</b>
4.3.1	SETTLING DOWN TO DO THE JOB .....	83
4.3.1.a	<i>What goes Through your Mind (cognition)</i> .....	83
4.3.1.b	<i>Mechanics or practicalities</i> .....	85
4.3.1.c	<i>How do you Mark or Grade a Student Paper (Assessment Artefact)?</i> .....	87
4.3.1.d	<i>Using the Marking and Grading Scheme</i> .....	89
4.3.1.e	<i>Environment</i> .....	92
4.3.2	PARAMETERS OF PRACTICE.....	94
4.3.2.a	<i>Being Reflective, Reflective Practitioner</i> .....	94
4.3.2.b	<i>Protecting Clinical Practice</i> .....	96
4.3.2.c	<i>Preparing new Staff</i> .....	98
4.3.2.d	<i>Marking and Grading Experience</i> .....	100
4.3.3	MESSINESS OF MARKING AND GRADING.....	103

4.3.3.a	<i>Arriving at Final Mark or Grade</i> .....	103
4.3.3.b	<i>Intuition – How to Calculate Mark or Grade</i> .....	105
4.3.3.c	<i>Question of Subjectivity</i> .....	108
4.3.3.d	<i>Moderation Processes</i> .....	109
4.3.3.e	<i>Reaching an Agreement on Mark or Grade</i> .....	112
4.3.3.f	<i>Marking/Grading vs Workload</i> .....	115
4.3.3.g	<i>Messiness of Marking and Grading</i> .....	119
4.3.4	KNOWING.....	124
4.3.4.a	<i>Assessment Artefact vs. the Requirements</i> .....	124
4.3.4.b	<i>Tacit Knowledge</i> .....	126
4.3.4.c	<i>Aware of Student Assessment Task/Knowing the Assessment</i> .....	127
4.3.5	MARKING AND GRADING BY-PRODUCTS.....	130
4.3.5.a	<i>Approach to Student Feedback</i> .....	130
4.3.5.b	<i>Being Consistent</i> .....	133
4.3.5.c	<i>Judgement and Justification</i> .....	134
4.3.6	SUMMARY .....	136
<b>CHAPTER 5 PROTOCOL ANALYSIS INTERVIEW METHOD .....</b>		<b>138</b>
<b>5.1 WHAT IS PROTOCOL ANALYSIS? .....</b>		<b>138</b>
<b>5.2 PROCEDURE .....</b>		<b>139</b>
5.2.1	CONDUCTING PA INTERVIEWS .....	140
5.2.2	ADVANTAGES .....	141
5.2.3	DISADVANTAGES.....	143
5.2.4	TRAINING .....	145
5.2.5	SELECTION OF PA INTERVIEWEES.....	145
5.2.6	MEMBER CHECKING .....	146
<b>5.3 SUMMARY .....</b>		<b>147</b>
<b>CHAPTER 6 PROTOCOL ANALYSIS FINDINGS .....</b>		<b>149</b>
<b>6.1 RESEARCH AIM.....</b>		<b>149</b>
6.1.1	RESEARCH OBJECTIVES: .....	150

6.1.2	RESEARCH QUESTIONS:	150
<b>6.2</b>	<b>INTERVIEWEE KEY 2</b>	<b>151</b>
<b>6.3</b>	<b>EXPLORATION OF FINDINGS</b>	<b>153</b>
6.3.1	ENVIRONMENTAL ELEMENTS:	154
6.3.1.a	<i>Anonymous Marking and Grading</i>	154
6.3.1.b	<i>Artefact Type</i>	155
6.3.1.c	<i>Impact of Workload on Decision Making</i>	157
6.3.1.d	<i>Learning Outcomes</i>	159
6.3.1.e	<i>Marking and Grading Environment</i>	161
6.3.2	STUDENT ASSESSMENT LITERACY:	162
6.3.2.a	<i>Academic Style</i>	162
6.3.2.b	<i>Analysis/Critical Analysis</i>	164
6.3.2.c	<i>Preparation, Practice, Purpose</i>	165
6.3.3	OPERATIONAL NECESSITIES	166
6.3.3.a	<i>Unknown or Unfamiliar Assessment</i>	167
6.3.3.b	<i>Conforming to Community of Practice</i>	168
6.3.3.c	<i>Generating a Mark or Grade</i>	169
6.3.3.d	<i>Marker or Grader Analysis of the Artefact Content</i>	172
6.3.3.e	<i>Marking and Grading Criteria Construction</i>	173
6.3.3.f	<i>Moderation Processes</i>	175
6.3.4	TANGIBLE MARKER OR GRADER ACTIONS	178
6.3.4.a	<i>Annotating Assessment Artefacts</i>	178
6.3.4.b	<i>Checking References</i>	180
6.3.4.c	<i>Constructing Feedback</i>	183
6.3.4.d	<i>Making Notes</i>	185
6.3.5	IMPLICIT KNOWLEDGE OR ACTIONS:	186
6.3.5.a	<i>Deciding on the Mark or Grade</i>	187
6.3.5.b	<i>Getting a Sense of or Overall Impressions</i>	190
6.3.5.c	<i>Interacting with the Artefact</i>	191

6.3.5.d	<i>Intuitive Device for Constructing Mark or Grade</i> .....	193
6.3.5.e	<i>Sense of Disappointment</i> .....	195
6.3.6	<i>'NEWBIE' ISSUES</i> .....	196
6.3.6.a	<i>Difficulties as an Early Career Marker or Grader</i> .....	196
6.3.6.b	<i>Learning about Marking and Grading</i> .....	198
6.3.6.c	<i>Preparation of New Staff</i> .....	200
6.3.6.d	<i>"Takes me Longer"</i> .....	201
6.3.7	<i>SUMMARY</i> .....	202
<b>CHAPTER 7 FUSION OF FINDINGS</b> .....		<b>203</b>
<b>7.1 RESEARCH SYNTHESIS</b> .....		<b>206</b>
7.1.1	<b>OBJECTIVE 1: EXAMINE ASSESSMENT BY EXPLORING MARKING AND GRADING PRACTICES</b> .....	207
	<i>Question 1 - What do lecturers believe they do when marking or grading? (From CI interviews)</i> .....	208
7.1.1.a	<i>Settling down to do the job</i> .....	208
7.1.1.b	<i>Parameters of Practice</i> .....	211
7.1.1.c	<i>Messiness of Marking or Grading</i> .....	213
7.1.1.d	<i>Knowing</i> .....	217
	<i>Question 2 - What thoughts or cognitive processes occur in the lecturer during the act of marking or grading? (From PA interviews)</i> .....	219
7.1.1.e	<i>Student Assessment Literacy</i> .....	219
7.1.1.f	<i>Operational Necessities</i> .....	220
7.1.1.g	<i>Tangible Marker or Grader Actions</i> .....	222
7.1.1.h	<i>Implicit knowledge or actions</i> .....	222
	<i>Question 3 - What exists to ensure equity or consistency in professional or inter professional marking or grading, this includes inter-rater reliability?</i> .....	224
7.1.1.i	<i>Operational Necessities</i> .....	224
7.1.1.j	<i>Tangible Marker or Grader Actions</i> .....	224
7.1.1.k	<i>Implicit Knowledge or Actions</i> .....	225
7.1.1.l	<i>Messiness of Marking or Grading</i> .....	225

SIGNIFICANT OUTCOMES .....	226
7.1.2 OBJECTIVE 2: IDENTIFY WHAT ARE ‘GOOD’ MARKING PRACTICES TO DEVELOP A CONCEPT OF ‘EXPERTISE’ IN MARKING AND GRADING.....	227
<i>Question 1 - What do lecturers believe they do when marking or grading? (From CI interviews)</i> .....	228
7.1.2.a Parameters of Practice .....	228
7.1.2.b Knowing .....	229
<i>Question 2 - What thoughts or cognitive processes occur in the lecturer during the act of     marking or grading? (From PA interviews).....</i>	230
7.1.2.c Operational Necessities .....	230
7.1.2.d Tangible Marker or Grader Actions .....	230
7.1.2.e Implicit Knowledge or Action .....	231
7.1.2.f ‘Newbie’ Issues .....	232
<i>Question 3 - What exists to ensure equity or consistency in professional or inter professional     marking or grading, this includes inter-rater reliability?.....</i>	232
7.1.2.g Tangible Marker or Grader Actions .....	232
SIGNIFICANT OUTCOMES .....	233
7.1.3 OBJECTIVE 3: EXPLORE NOVICE LECTURERS THOUGHTS ON MARKING OR GRADING.....	234
<i>Question 1 - What do lecturers believe they do when marking or grading? (From CI interviews)</i> .....	235
7.1.3.a Settling Down to do the Job.....	235
7.1.3.b Messiness of Marking or Grading.....	236
<i>Question 2 - What thoughts or cognitive processes occur in the lecturer during the act of     marking or grading? (From PA interviews).....</i>	237
7.1.3.c Student Assessment Literacy .....	237
7.1.3.d Operational Necessities .....	237
7.1.3.e Tangible Marker or Grader Actions .....	238
7.1.3.f Implicit Knowledge or Actions .....	239
7.1.3.g Newbie Issues .....	239

<i>Question 3 - What exists to ensure equity or consistency in professional or inter professional marking or grading, this includes inter-rater reliability?</i> .....	241
7.1.3.h <i>Operational Necessities</i> .....	241
7.1.3.i <i>Implicit Knowledge or Actions</i> .....	241
7.1.3.j <i>Messiness of Marking or Grading</i> .....	242
SIGNIFICANT OUTCOMES .....	243
7.1.4    OBJECTIVE 4: EXPLORE COGNITIVE PROCESSES AND EXTRANEOUS INFLUENCES ON MARKING OR GRADING PRACTICE	244
<i>Question 1 - What do lecturers believe they do when marking or grading? (From CI interviews)</i> .....	245
7.1.4.a <i>Settling down to do the job</i> .....	245
7.1.4.b <i>Parameters of Practice</i> .....	248
7.1.4.c <i>Messiness of Marking or Grading</i> .....	251
<i>Question 2 - What thoughts or cognitive processes occur in the lecturer during the act of marking or grading? (From PA interviews)</i> .....	252
7.1.4.d <i>Operational Necessities</i> .....	252
7.1.4.e <i>Tangible Marker or Grader Actions</i> .....	253
7.1.4.f <i>Implicit Knowledge or Actions</i> .....	254
<i>Question 3 - What exists to ensure equity or consistency in professional or inter professional marking or grading, this includes inter-rater reliability?</i> .....	255
7.1.4.g <i>Environmental Elements</i> .....	255
7.1.4.h <i>Operational Necessities</i> .....	257
7.1.4.i <i>Messiness of Marking or Grading</i> .....	258
SIGNIFICANT OUTCOMES: .....	261
7.1.5    SUMMARY .....	262
SUMMARY OF RESEARCH OUTCOMES (RO).....	263
<b>CHAPTER 8    CREATIVE SOLUTIONS</b> .....	<b>266</b>
<b>8.1    MODEST BEGINNINGS</b> .....	<b>266</b>
<b>8.2    REFLEXIVITY</b> .....	<b>268</b>
8.2.1    THESIS CHAPTER SUMMARY .....	270

<b>8.3</b>	<b>LIMITATIONS.....</b>	<b>273</b>
<b>8.4</b>	<b>MAIN FINDINGS .....</b>	<b>274</b>
8.4.1	“VENEER OF GOODNESS” .....	274
8.4.2	“THAT IS WHY I JUST CAN’T LOOK INTO THEIR FACES” .....	281
<b>8.5</b>	<b>CONTRIBUTION TO CONTEMPORARY KNOWLEDGE .....</b>	<b>284</b>
<b>8.6</b>	<b>FROM NOVICE TO EXPERTISE: JOURNEY TO EXPERT PRACTICE .....</b>	<b>289</b>
8.6.1	DEVELOPING EXPERTISE IN MARKING AND GRADING PRACTICE: .....	290
<b>8.7</b>	<b>SUMMARY .....</b>	<b>294</b>
<b>8.8</b>	<b>RECOMMENDATIONS.....</b>	<b>295</b>
	<b>EPILOGUE.....</b>	<b>297</b>
	<b>REFERENCE LIST.....</b>	<b>298</b>
	<b>APPENDICES .....</b>	<b>319</b>
1.	<i>Ethics Approval.....</i>	<i>319</i>
2.	<i>Interviewee Information Sheet.....</i>	<i>324</i>
3.	<i>Sampling Frame .....</i>	<i>327</i>
4.	<i>Recruitment Protocol .....</i>	<i>328</i>
5.	<i>Consent Forms.....</i>	<i>329</i>
6.	<i>CI Protocol .....</i>	<i>331</i>
7.	<i>PA Interview Schedule .....</i>	<i>334</i>
8.	<i>PA Training.....</i>	<i>336</i>

# TABLE OF TABLES

Table 1 Glossary of Terms.....	xvi
Table 2 Comparison of Skill Acquisition Models .....	22
Table 3 CI Interviewees .....	79
Table 4 PA Interviewees .....	151
Table 5 Research Objectives and Questions.....	205
Table 6 Objective 1 .....	207
Table 7 Objective 2 .....	227
Table 8 Objective 3 .....	234
Table 9 Objective 4 .....	244
Table 10 Summary of research outcomes .....	265



# TABLE OF FIGURES

Figure 1 Competence quadrant .....	26
Figure 2 Original Static Model.....	58
Figure 3 Original Major Themes and Sub Themes .....	60
Figure 4 Marking and Grading Processes.....	72
Figure 5 CI Final Major Themes and Sub Themes .....	81
Figure 6 CI Major Themes .....	83
Figure 7 CI Major Theme .....	94
Figure 8 CI Major Themes .....	103
Figure 9 CI Major Theme .....	124
Figure 10 CI Major Theme .....	130
Figure 11 PA Initial Coding .....	152
Figure 12 PA Major Theme .....	154
Figure 13 PA Major Theme .....	162
Figure 14 PA Major Theme .....	166
Figure 15 Major Theme.....	178
Figure 16 PA Major Theme .....	186
Figure 17 PA Major Theme .....	196
Figure 18 Attributes of expertise in marking and grading.....	290
Figure 19 Fundamental requirements for novice evaluators .....	291

# GLOSSARY OF TERMS

Below is a list of words and phrases which are used in everyday academic practice and academic discourse. However, it is easy for commonly held expressions to have meanings which are as diverse as the people interpreting them. Therefore it is pertinent that an explanation is provided of the meaning of these words in this study.

<b>Assessment/ Assessment artefact</b>	Refers to what has been produced by the student to demonstrate their level of learning. In this research the assessment artefact will be a written essay and will be subject to being marked or graded as a summative assessment.
<b>Marking</b>	Will indicate the activity undertaken by lecturers to validate student attempts at an assessment and will always denote evaluating written essays. (Used interchangeably with grading - to capture practices beyond the UK)
<b>Grading</b>	Will indicate the activity undertaken by lecturers to validate student attempts at an assessment and will always denote evaluating written essays. (Used interchangeably with grading - to capture practices beyond the UK)
<b>Lecturer/ Evaluator</b>	Individuals employed in higher education in the teaching and assessment of students via marking and grading.
<b>Expertise</b>	The possession of more knowledge and or skill in a particular area or specialty than could be expected to be known by the general population.
<b>Student</b>	An individual studying on a named programme to obtain a qualification or certificate.
<b>Practitioner</b>	A lecturer immersed in student assessment through marking and grading student work
<b>Practitioner- Researcher</b>	An individual conducting research into practices which they are engaged in, generally within their own work/life context.
<b>Interviewee</b>	Lecturer participating in the research by completing either a Protocol Analysis interview or Cognitive Interview.
<b>Rubric</b>	A guide for scoring achievement in the assessment artefact. Uses criteria, generic or specific which may be split into a number of elements, in the form of qualitative statements, signifying achievement along a continuum.

Table 1 Glossary of Terms

# PROLOGUE

To enable you, the reader, to make sense of all of which is to follow it is important to do two things, introduce myself and the impetus for this work.

I am a female lecturer, with 16+ years' teaching experience in Higher Education. This career has followed a previous 17 years in clinical practice in the National Health Service, where even back then we were being encouraged to deliver care which was evidence based. I arrived in Higher Education having completed a Postgraduate Diploma [PgDip] in Health Professional Education. I was keen to learn more about all facets of academic practice, of which marking and grading student work was one such aspect. I was surprised with my initiation to this part of the lecturer role; it appeared to lack the rigour I had associated with my clinical training and practice.

I learned to mark by doing, and then having the privilege of comparing my marks with one of my most senior colleagues in the department. I remember feeling embarrassed and inadequate as I found it difficult to interpret the learning outcomes and conceptualise whether the essays submitted for marking and grading matched what was required. I reflected on the module I had completed on 'assessment' in my PgDip, there seemed to be little preparation for the role I had taken on. This was certainly a limitation in the module, as no attention had been given to the marking and grading aspect of 'assessment'. The focus was located with method, exploring which would be most appropriate to gauge student knowledge development i.e. would the use of an examination be more effective than requiring the student to draft an essay.

Early in my second academic year of lecturing, still a novice in relation to marking and grading, I became an active member of a Faculty wide working group. The remit of this working group was to develop generic marking and grading criteria or rubric for student assessment, to be used by all academics when evaluating assessment artefacts. Over the following 8 years I worked closely on supporting the continued review and development of the criteria

and led attempts at integrating and embedding the criteria into departmental assessment practices within and beyond the Faculty.

I therefore do not enter this PhD to be immersed in a totally new topic. Instead I bring to the subject a wealth of experience in developing tools to support those who mark and grade. However, from reflecting on my own experiences there is something that needs to happen even before a lecturer can think of applying criteria, this is knowing the know-how of marking and grading. At the end of the research process I am in a position to respond to this notion of knowing the know-how. My research reveals the existence of a concept of a know-how of marking and grading, and suggests potential ways in which this knowing can be articulated and thus learned by novice lecturers.

***“When you come to the edge of all the light you know, and are about to step off into the darkness of the unknown, faith is knowing one of two things will happen: There will be something solid to stand on, or you will be taught how to fly.”***

Barbara J. Winter

# INTRODUCTION

This initial introduction to the thesis is presented to facilitate navigation amongst the content. The purpose of this thesis is to demystify marking and grading, and in doing so, identify expert practices. However, it is important to explore the drivers which underpin my approach to data collection and data analysis. These drivers are presented in the prologue providing context to my history of, and connections with marking and grading. The introduction then goes on to propose a guide to the coming content whilst latterly it begins to set the scene by presenting the background to the research.

## THESIS STRUCTURE

This thesis is structured in the following way; Chapter 1 is concerned with a review of the literature and exploration of the notion of expert practice. Chapter 2 then discusses the underpinning philosophical orientation and considers the methodological decisions taken in light of this. It is in this chapter where the research aim, research objective and subsequent research questions are developed and refined following close attention to current evidence on marking and grading and my espoused ontology and epistemology. This chapter does not include detail of the two methods selected to gather data from the research interviewees. Instead, Chapter 3 gives details of the first interview method, Cognitive Interviewing. Directly following this comes Chapter 4 which delivers the findings from this method. Chapter 5 then is given over to an explanation of the second method of data collection utilised, Protocol Analysis. Chapter 6 then reports the findings from this method. The rationale for this structure lies with ensuring close proximity between reading about the method and being able to review its application and subsequent findings whilst retaining some familiarity with the procedure. Chapter 7 becomes the synthesis of the findings and it achieves this by close attention to the research objectives. Each objective has a set of 3 common research questions applied to it in order to uncover ideas relevant to the topic of that objective. The research objectives and research questions can be located in Chapter 2. Examination of interviewee transcripts reveal the existence of duplicate themes which arise from the two

interview methods. This research therefore exposes issues which are of great significance to those marking and grading and thus warrant further investigation. The final chapter (Chapter 8) is one which draws together the findings, explores any limitations of the data collection methods and approaches to data analysis whilst reporting conclusions. It is here that the opportunity is taken to present my contributions to existing knowledge on marking and grading in HE. I present a model which has the potential to be utilised to support those new to marking and grading along with proposing ideas for further research.

At the outset it is essential to highlight that there are times when this thesis is written in the first person, where it seems to be appropriate. The research methodology used in this thesis has taken a post positivist stance, it utilises phenomenology and involves the exploration of lived experiences of lecturers as study participants or interviewees. This in-depth investigation enabled identification of participants' specific beliefs, values and traditions relating to marking and grading student written assessment artefacts. I contend that as the researcher my starting point will inevitably be influenced by preconceptions or biases; in using the first person I am making clear how interpretations and meanings have been derived by making the researcher 'visible' in the research. Webb (1992) asserted this to be entirely appropriate, and in many cases it is more truthful to use the first person rather than being:

*“... obliterated as active agents in the construction of knowledge”.*

(page 748)

## **SETTING THE SCENE**

A brief review of the Higher Education landscape acts as a catalyst for establishing the significance of exploring marking and grading as a subject, by presenting background evidence of its relevance to academic practice.

Traditionally university students came from environments where entry to higher education is the norm rather than being an unusual occurrence. Most commonly these students have parents who have continued their education beyond formal school years, and are found to have good socio-economic backgrounds (Office for National Statistics 2016), with residency in non-deprived areas. A strategy to widen access and participation in higher education has been encouraged, during the preceding 15 years, by successive governments. Since 2001 there has been an increased allocation of funding via the Department for Innovation, Universities and Skills, and Higher Education Funding Council with its success or otherwise being reviewed in 2007 - 2009 (House of Commons 2009). The purpose of this strategy was to increase numbers of students from non-traditional backgrounds entering higher education, most significantly young males from lower socio-economic backgrounds, including those from white ethnic backgrounds residing in deprived areas.

An increased educational background potentially increases the social standing of graduates by opening up access to jobs commanding higher salaries than their parents, for whom education terminated at the school leaving age. Education is then valued as a commodity, for what it can do, for the individual, and therefore for society, by contributing to and increasing its overall wealth. Dearing (1997) asserts that students should exit HE being able to be responsible for their own continued learning and development, contributing to a learning society. This last point is interesting as it highlights the value placed on education and educational achievement, with it being seen as a way of developing individual and societal prosperity. This function of education is seemingly held in higher esteem than the act of learning itself, and the acquisition of knowledge, which then has limited value as an enterprise, in itself (Füredi 2009).

Education is thus viewed as a valuable commodity, and accordingly the way in which that educational achievement is measured takes on a sense of greater importance. Reflecting on these changes identifies a growing heterogeneity within the HE student population impacting on approaches to teaching, concepts of learning and modes of assessment of educational achievement. Widening participation has resulted in a student population which has a varying number of needs for learning and physical adjustments. It is evident that the increase in student numbers with the massification of HE has brought with it the notion of an inevitable reduction in academic standards (Leathwood 2005). This perspective only continues to reinforce the societal divide in those who seem to warrant access to the privilege of an education and those who have been allowed to participate. Thus the student population can be challenging in terms of working with many different requirements which may be protected through the Equality Act (2010) and the preceding Special Educational Needs and Disability Act ([SENDA] 2001). Further and Higher Education institutions have a responsibility to not 'victimise a disabled person':

- a) in arrangements for deciding upon whom to confer a qualification
- b) a duty to make reasonable adjustments applies to the responsible body of such an institution

(Equality Act (*Equality Act* 2010) c15, Part 6, Chp 2, 90 – 94 page 58)

Higher Education has no immunity from the social context and current approaches to inclusivity and diversity.

Such diversity in the student population is not always matched by diversity in the range of lecturers and professors responsible for teaching, learning and assessment. The constitution of the academy is largely representative of the majority population, with as few as 1% of ethnic minority economic academics being born in the UK (Blackaby and Frank 2000). The result is a homogenous body, with a limit in their knowledge on inclusivity and diversity, being responsible for conferring awards on a heterogeneous study group. Lecturers may need preparation for reviewing assessment artefacts which, for one reason or another, are presented in a non-traditional format,



constructed with support of assistive technologies. The imperative for such consideration becomes evident when exploring the literature. A study by Fuller et al. (2004) reported difficulties experienced by disabled students with participation in assessment which could be administered via modes including written approaches such as examination and coursework, or oral presentations. Moreover a study conducted in 2010 (Madriaga et al.) using a matched sample, conveyed participants experience by identifying the similarities faced by disabled and non-disabled students, with the many assessment approaches.

Assessment usually encompasses a number of methods to support the student in presenting their knowledge for a completed unit of learning. Knight (1995) asserts that there could be as many as 50 different methods which could be used to gather sufficient evidence of student achievement. The more commonly used assessment methods are often one of the following: essays, presentations, examinations, posters and portfolios, with the output from these methods being referred to in this thesis as assessment artefacts. Any one of this range of assessment artefacts can be produced by the student for evaluation purposes. It is this evaluation that generates feedback which is used for formative or summative purposes. Black and Williams (2009) provide a definition of formative assessment as that which is *"concerned with the creation of, and capitalization upon, 'moments of contingency' in instruction for the purpose of the regulation of learning processes"* (p10). Whereas summative assessments *"provide ways of eliciting evidence of student achievement .... can prompt feedback that moves learning forward... communicating criteria for success"* (p8). Feedback on summative assessment could consist of solely a mark or could be a mark and qualitative text which could provide the student an insight into the decisions made by the lecturer when interpreting and evaluating the student assessment artefact. Summative assessment does emphasise achievement of the learning outcomes (Carless 2006), valuing the product of learning rather than the actual learning journey itself. It does this by making a contribution to the final award for student study.

Retaining quality in assessment remains the end game, being balanced against a requirement to make 'reasonable adjustment', to ensure all assessments are accessible by all students. This may necessitate conceptualising a different approach to the assessment of student knowledge, whilst not reducing standards nor overtly advantaging disabled students (Ashworth et al. 2010). The difficulty here arises with deciding on the effectiveness of dissimilar assessment approaches being used to test for acquisition and achievement of the set learning outcomes, whilst providing assurance of an equitable level of difficulty (Crisp and Novaković 2009). This approach must be taken if there is to be continued assurance of maintaining standards in HE qualifications (QAA 2012; Leathwood 2005).

Students who participate in Higher Education expect to exit their studies with a qualification or certification that authenticates the acquisition of new knowledge or development of practical or technical skills. Learning is monitored, recorded and measured against set outcomes or requirements for an award by review and evaluation of artefacts produced for assessment. Standards of achievement are specified by set outcomes for multiple assessment artefacts; evaluation of these facilitate the quantification of attainment on completion of requirements for an award. Achievement of these standards is identified through marking and grading, an activity undertaken by lecturing staff, which facilitates the appropriate award being conferred upon each student on completion of the unit or course of study. The accuracy of artefact evaluation becomes important when considering the potential to deliver spurious information on academic ability, such as under or overachievement, to students. Inaccurate judgements of the value of an artefact have potential to impact on the award or denial of certification of academic accomplishment. Exploring current approaches to marking and grading of student written assessment artefacts thus presents itself as a worthwhile endeavour.

The Quality Assurance Agency for HE (QAA) add weight to this argument following close inspection of a document constructed to provide support to new academic staff. The document 'Understanding assessment: its role in

safeguarding academic standards and quality in higher education' (QAA 2012) is devoted entirely to assessment. It contains much information on the implications of decisions that lecturers may make regarding the assessment type to be deployed, from the design through to its delivery. It explores assessment through module and programme review, developing new academics understanding of assessment processes. There remains an absence of information or instruction on how a lecturer might undertake the process of marking and grading of any assessment artefact submissions. This is an oversight which has the potential to influence the development of new academics.

The aim of the research reported here was to discover what occurs during the process of marking and grading assessment artefacts, and how lecturers carry out this component of their role. The focus was on identifying thoughts and actions which lead to the evaluation and generation of a mark and grade for an assessment artefact.

# Chapter 1 LITERATURE REVIEW

*He who knows not, and knows not that he knows not, is a fool - shun him,  
He who knows not, and knows that he knows not is ignorant - teach him,  
He who knows, and knows not that he knows, is asleep - wake him,  
But he who knows, and knows that he knows, is a wise man - follow him.*

Anon.

This chapter is given over to exploring contemporary literature relevant to marking and grading in HE, and begins to focus on components that enables a contextualising of the topic. The literature review has been constructed through an iterative approach, this maintains a close connection with the purpose of the research, whilst exploring what is currently known about lecturers marking and grading in HE.

When commencing any research study it is important to identify what exists and what is already know about the topic prior to taking a narrower focus which is specifically relevant the topic or area of interest. Therefore a review of the literature commenced with wide scoping activity utilising terms for searching which I considered were relevant to the topic of lecturers marking. Much of this early literature identified has been discussed in the section 'Setting the Scene' (p xx- xxiv) which succeeds in placing the rationale for this study into context of contemporary thought about marking and grading and its importance.

Databases and publisher websites were used to begin to source appropriate articles which consisted of research papers and opinion papers on topics related to marking and grading. The following is a list of the majority but not exhaustive list of external sources consulted:

British Humanities Index/Proquest  
BUBL Link; Education  
Cambridge Uni Press Online Journals  
EBSCO Electronic Journals  
Education -Line  
Emerald-Journals  
ERIC  
HEFCE Publications  
HEA Resources  
IDP Database of Research Internat Educ.  
JORUM

Research into HE Abstracts  
Science Direct  
Taylor Francis  
UK Official Pub Online  
Web of Knowledge  
Wiley Online

Early parts of this thesis and research process were focussed on identifying avenues to explore for development the research questions. An approach to reviewing the literature was taken which sought to contain the scope of the review to those areas (Boell and Cecez-Kecmanovic 2014). These authors determine how a literature review can be undertaken which has at its core a hermeneutic approach with a desire to uncover literature which is not pre-determined as it would be in a systematic review. Instead they advocate a process which facilitates movement between different texts in an iterative way as understanding around the topic increased. This process facilitates following avenues which emerge during the course of that reading by undertaking further searches reading literature referred to in texts or indeed identifying newly published sources or those which may have been initially overlooked.

Therefore the literature review presented here is the product of such a process and not one which was conceived in a chronological order, instead it responds to the research questions and to findings following data collection. It does this whilst not relinquishing opportunities to further enhance a demonstration of depth of reading not solely breadth.

This literature review is divided into two. The first is concerned with exploring concrete elements of the marking and grading process whilst the second section explores more conceptual themes identified as being related to the practice of marking and grading. The terms marking and grading are used here interchangeably to signify the activities undertaken by lecturers, taking into account the differing scales used by institutions to signify student achievement. The first section therefore explores the following areas: marking or grading criteria and rubrics; normative and or criterion referencing; completing marking and grading as an activity; assessment of literacy of staff and students. The second section explores the following

conceptual themes: tacit beliefs in relation to marking and grading; expertness including an examination of the notion of expert practice; the relevance of communities of practice for achieving marking and grading and ending with an examination of what it means to learn a new skill.

For the purpose of this thesis, all assessment submissions being considered are written essays and are referred to as assessment artefacts. The use of essays for this study will facilitate lecturers in talking about how they reach their evaluation decisions. This is different to that which would be achieved when compared to narrower assessment tasks such as multiple choice paper assessment where answers are either correct or incorrect with limited margin for error (Yorke et al. 2002).

## SECTION 1

### 1.1.1 MARKING OR GRADING CRITERIA, AND RUBRICS

Marking or grading criteria and rubrics are developed to facilitate the review of student assessment artefacts submitted for an evaluation of performance against requirements on completion of a module, unit of study or course. These rose in popularity from the mid 1990's as a way of reducing the possibility of an arbitrary evaluation of an assessment artefact, ensuring a more consistent, structured and reproducible approach to marking and grading (Sadler 2009c). Milligan (1996) maintains that the criteria which are to be used for judging student achievement should be clarified and available in the public domain. Evaluations seemingly based on professional judgement alone rely on tacit knowledge with the standards required for successful achievement remaining hidden from students and colleagues (Brooks 2012) and thus can be viewed as an unethical practice. Students are unable to prepare effectively for such types of assessment because the standards that they are required to meet remain unknown (Milligan 1996). The use of published marking and grading criteria thus can be viewed as an ethical approach to assessment practice.

When marking or grading, lecturers assess the perceived 'worth' or value of the submitted artefact. The final summation or evaluation of the submitted

assessment artefact may be reflected in a textual account of student achievement via qualitative comments or statements. Qualitative statements are frequently accompanied by presenting student achievement using a numerical mark, with a common approach being to use both (Bridges et al. 1999). Student achievement can also be illustrated by applying the use of traditional grades using letters A – E (Yorke 2011) and grades formulated on a numerical base, with traditionally an ‘A’ grade being worth 70 - 100%, whilst an ‘E’ being 39% or less. This system has been commonly used in education within the United Kingdom. Another method of illustrating student assessment is to use the same descriptors which are used to award degree classification. These would equate to students being awarded a system of marks which clearly placed the student in an award category of 1:1 (first class honours), 2:1 (upper second class honours) 2:2 (second class, second division), 3 (third class with or without honours) (Yorke et al. 2002). Or indeed students’ achievement is couched in terms of fail, pass, merit, or distinction. In many cases what students receive is a blend of any of these systems, with the numbers often being required as a route to interpret these descriptive achievements (Sadler 2005). Use of such signifiers to describe student achievement is a common approach, although Yorke (2011) finds this to be a contentious issue. It presents a student with a quantified symbol which belies the subjectivity of marking and grading, objectifying it by creating a level of certainty in the mark awarded.

Marking and grading can occur by reviewing student work in light of stated standards expressed as criteria or rubrics. Sadler (2009b) highlights the utility of sets of text based statements of achievement which are grouped together to form matrices which can be used by lecturers to evaluate the value of an artefact submitted for assessment. The assumption is that required student outcomes, in the guise of objectives or standards, can be articulated effectively to enable lecturers to be able to apply these without difficulty to the assessment artefact. Assessment completed using criterion referencing is achieved via secure and agreed external reference points are purported to facilitate a means of transparency for anchoring judgements made about an assessment artefact (Brooks 2012).

Jonsson and Svingby (2007) investigated the reliability, validity and educational consequences of rubrics through the review of 75 empirical studies on rubrics. They established that under certain conditions the use of a rubric or marking and grading matrices had the effect of improving the reliable scoring of assessments when rubrics were analytical, topic specific and included evaluator training, and or the provision of exemplars. Their review reiterated that it was more probable that assessment with a rubric produced a much more reliable assessment of performance than assessment without. Rubric use was also shown to improve intra and interrater reliability in the majority of studies they reviewed (which used this as a measure). There remains a level of doubt of the validity of such rubrics related to the difficulties in establishing the ability of these tools to measure what they set out or intended to measure (Popham 1997; Hack 2015). Sadler (2009c) is derisive of the ability of criteria based analytical systems of assessment to deliver objective and therefore ethical evaluation of assessment artefacts. Studies reviewed by Jonsson and Svingby (2007) presented limited evidence of the validity of rubrics, with the concept of construct validity being commonly used to encompass different perceptions of validity, for example content validity, generalizability, and consequential validity. These relate to the ability of the assessment to attain a representative measure of knowledge and skills portrayed by the student; for this rubric to be used to facilitate interpretation across a number of groups, assessment tasks and time periods. Consequential validity pertains to the ethical perspective of the rubric for considering the implications of interpretation in relation to identification of student achievement.

Research identified that despite the presence of criteria for assessment, without a careful and shared interpretation by lecturers it was difficult to ensure a consistent approach to artefact evaluation (Saunders and Davis 1998). Such an approach would incorporate specifics of how to apply the criteria, and cultivate increased knowledge and ownership which is more easily established when lecturers have the opportunity to be involved in their development. At the point of implementation for assessment, participating in negotiating with colleagues in relation to use in their current context will



increase consistency. Hand and Clewes (2000) identified that their participants would not always use or know or apply guidelines for artefact construction or marking and grading criteria to carry out their assessment or evaluation, instead appearing to rely on their tacit knowledge and standards. Such an approach reduces the possibility of consistent application and shared interpretation between the ranges of evaluators involved in the assessment of a single essay across a student cohort.

One of the advantages of using a set of marking or grading criteria, or rubrics is the consistent evaluation of specific elements in the assessment artefact. Criteria are developed to guide the lecturer in exploring the artefact in a structured way, evaluating set areas before coming to a global or holistic estimation of its value (Saunders and Davis 1998). There is potential for different lecturers to place different emphases on different elements of the criteria, explaining how different marks can be awarded by different lecturers for the same assessment artefact. It seems that lecturers not only apply different weightings to different criterion, but also use personal criterion to evaluate the assessment artefact (Hand and Clewes 2000). Inevitably, this works against the ability to perform equitable and comparable evaluation between lectures, leading to inconsistent student outcomes being achieved. These authors suggest this promotion or demotion of criteria weighting leads to operating a system of primary or secondary criteria by individual lecturers. Such a position means that these accents on criterion may not be visible or indeed remain hidden from students and colleagues.

Ashworth et al. (2010) revealed how lecturers attempted to reconcile their approach to assessment and assessment standards in light of the perceived ability of students with an impairment or disability. Whilst they utilised the institutional criteria for evaluation of the assessment artefact, at the level the student was entered to study at, they felt they needed to adapt their interpretation of what was required in order to demonstrate achievement. To prevent lecturers seeing this as a lowering of standards, participants '*recast student achievement as different rather than inferior*' (p 218). These findings demonstrate that not all criterion by which student work is judged can be

explicitly expressed and instead remain obscured. Shay (2004) questions the trustworthiness of academic evaluations of assessment artefacts by exploring the foundations upon which interpretations are made. It was clear different lecturers would give priority to different elements of their evaluator functions, this may be relative to the criteria for assessment or it may be related to their advantaged position as supervisor. Knowing the student, working with them in their development of the assessment artefact impacts upon the lecturer's perspective of the finished product. As Shay (2004) highlights, any evaluation of the artefact is made from an insider position, being furnished with knowledge that is unavailable to other evaluators such as those completing moderation processes.

Saunders and Davies (1998) further highlights why a contemporary approach to criteria review, and discussions on their utilisation is warranted; their findings suggest that lecturer's understanding and application changes over time. This point is further supported by Sadler (2009c) with the suggestion that criteria are interpreted differently by different lecturers and that it remains possible for the same lecturer to make different interpretations in different contexts. This could be relative to the purpose of the mark or grade which is prioritised by individual evaluators or indeed the community of practice. Interpretation of achievement is viewed as evidence of its value, as judged against conceptions of what constitutes good, bad or indifferent via collegiate consensus, whilst simultaneously being evidence of approval or otherwise of the artefact (Butler Shay 2004). This presents further evidence of what Yorke (2011) terms as a 'fallacy' in relation to the ability to measure achievement on a numerical or alpha numerical scale.

These issues are examined in this thesis when exploring the nature of marking and grading through application of the proposed research methods, and subsequent analysis of data (Chapters 4, 6, and 7).

### 1.1.2 NORMATIVE AND OR CRITERION REFERENCING

Sadler (2009b) discusses the utility of normative referencing against a backdrop of the possibility of criterion referencing. The first facilitates the

awarding of grades following inter-student comparison amongst their group of peers within a single study cohort. They contend that individual students are recruited into a cohort, of which they become a member, and cannot control the abilities and levels of achievement of these peers. They could be fortunate enough to be recruited into a low achieving cohort and thus their achievement be elevated above a level where it would be amongst a higher achieving cohort. Each student's level of achievement is thus reliant upon members of their peer group, which is not representative of a fair or equal process. When criterion referenced assessment is in operation it is thought to increase transparency, with students being required to meet specific criterion (for themselves) rather than competing with other students in a normative referenced system (Leathwood 2005).

Dubey and Geanakoplos (2010) explore varying reasons for our fascination with grades and marks. Use of various structures can provide external observers and students with particular impressions of the individual module or unit of study, in relation to the success of enrolled students. The authors undertake an analysis in which they compare what they denote as either absolute (where marks denoted by a percentage were provided) or relative approach of course grade bands where student performance is marked or graded on a curve and is related to the performance of others in the group. These two strategies fit well with the ideas of criterion and normative referencing. These authors contend that careful selection of the right approach can be used to influence student study behaviour, increasing competition between candidates, thereby increasing attainment.

As alluded to by Dubey and Geanakoplos (2010) there is the split between exclusively using explicit criteria to make a judgement about student achievement, rather than using normative referencing which adjusts student achievement depending upon marks or grades of others in their cohort. This latter concept is where all of the student submissions are placed in rank order. To do this the marker would need to decide which piece of work they regard as the 'average' level of achievement within the student cohort. Next each piece of work is compared to the one which was seen as the average

or 'norm' within that student group. Larrington and Roger (2002) contend that normative referencing systems can inflate or deflate the value of work by a perceived need to not over award high i.e. 1<sup>st</sup> class honours degrees.

Criteria are used by lecturers to compare an assessment artefact against, to establish the level of achievement. However Saunders and Davies (1998) challenges this simple explanation, suggesting that lecturers have an inability to apply criteria without interpreting them through previous experiences of evaluating assessment artefacts, thus the criterion are themselves norm referenced. Brooks (2012) also doubts that consistent application of published criteria as being the only means by which judgements of the value of an assessment artefact are reached. Lecturers' individual interpretations of the stated criteria impact upon the final decisions regarding evaluation of the assessment artefact, which are norm referenced.

One benefit of using stated criteria against which to assess students is the ability for those being assessed to be aware of the requirements for the assessment task, facilitating self-assessment during preparation and prior to submission of their assessment artefact. Where no criteria exists Sadler (2009b) suggests that students are then prey to an assessment which is related to individual lecturers' own set of standards, preferences or peculiarities whereby the student is not party to the requirements for assessment success.

It is evident, that to facilitate consistent evaluation, lecturers need to employ the same interpretive framework which is referenced in the same steadfast way. This is further complicated by the concept of evaluation. This could be executed by applying each criteria or norm standard individually to the assessment artefact which is named an analytical application of criteria by Saunders and Davies (1998). The other option is that lecturers could opt to take a global perspective of the artefact, taking into account the normative or criteria standards.

### 1.1.3 COMPLETING MARKING AND GRADING AS AN ACTIVITY

Student success or failure is determined by marking and grading which is carried out to evaluate the artefact submitted for assessment purposes. Therefore assurance in the robustness of evaluation processes undertaken is required, including bringing attention to any issues with validity and reliability. On examination of the literature, there is a variety of information and discussion about several facets of assessment including the evaluation of standards, Hornby (2003) explores this in relation to different academic disciplines and the use of grading criteria. It is clear that in recent years there has been increasing discussion regarding the act of marking and grading in HE (Baume et al. 2004; Boyd and Bloxham 2013). It is within this context that this research identified, explored and examined the factors that affect marking and grading of student work.

Marking and grading activity does present issues of validity and reliability for those submitting an artefact for assessment and for lecturers completing the evaluation. Granger et al (2008) indicates in their findings that the processes utilised by lecturers to arrive at their final evaluation are poorly understood and requires further study. Whatever method being used to judge the quality of the submitted artefact, it needs to consistently be able to be applied by a number of different lecturers, in different contexts and across a range of assessment artefacts. This is notwithstanding a need to ensure that which are criteria selected to evaluate the artefact has the power to do so, and that the tool applied needs to acquiesce to powers of validity and reliability.

### 1.1.4 ASSESSMENT LITERACY OF STUDENTS AND STAFF

Sadler (2009b; 2009d) states that when presenting marking and grading criteria or rubrics to students, the assumption is that these will be able to effectively communicate assessment expectations, such that students will be able to identify what they need to do to demonstrate understanding of and achievement at the task. Hand and Clewes (2000) would agree, seeing the use of criteria as one way of reducing the discrepancies in marking, making the decisions made become more transparent for staff and students. Rust et

al. (2003) demonstrated, students considered that the utility of marking and grading criteria was limited if there was no opportunity to discuss and demonstrate its application, understanding thereby remained a nebulous concept.

The use of a rubric also ensures that the requirements of the assessment are made more transparent and explicit, having possible implications on improving student learning (Jonsson and Svingby 2007). Once students are able to understand the assessment criteria it can be instrumental in bringing about changes in students' perceptions of their learning. However, Saunders and Davies (1998) asserts that interpretation of criteria for evaluation differs between individual lecturers and that it can then be assumed that students themselves will also hold different interpretations. To facilitate utilisation of criteria by students to make improvements to their assessment artefacts it is necessary for both lecturers and students to hold the same understanding and interpretation. This would be more easily achieved in the face of explicit criteria rather than via normative referencing. Exploring marking and grading criteria with students needs to be a joint and purposeful act which takes place at the earliest opportunity, prior to construction of the assessment artefact.

To perform well at any assessment task the student needs to understand what is required of them not only through understanding what the finished artefact will look like, but also how to define the learning outcomes so that these can be portrayed accurately. Bloxham and West (2004) define this process as a social practice which requires the integration of tacit knowledge and development of particular forms of written and verbal expression. These researchers introduced an element of peer marking, reporting on aspects of developing student understanding of assessment and its processes including application of assessment criteria, leading to improvements in their assessment literacy. Sadler (2009d) defines this as developing student assessment "*connoisseurship*", students being able to identify and apply explicit and latent criteria in constructing and evaluating their assessment artefacts. In order to operate effectively students need to learn a new

academic literacy which will enable them to complete their assessments more successfully. Without access to and understanding of the new vocabulary, students cannot all participate on an equal footing.

Student assessment literacy is acknowledged as an important component in successful participation in an end of study evaluation of learning, via summative assessment. Lea and Stierer (1998) highlighted that students could not engage effectively with assessment feedback as they lacked the appropriate 'academic literacy'. It is known that engagement with feedback on assessment activity is a way of improving future performance, but this is thus ineffective if students are unable to fruitfully engage with it, because of a lack of understanding. In order to engage students as partners, as suggested by Boud and Falchikov (2006), students need to be developed in terms of their understanding and approaches to assessment. This will prevent them being the subjects of assessment, and become 'active' rather than passive agents in this process. These and other authors (Bloxham and West 2004; Smith et al. 2013) have tried to increase students understanding of the way that their assessment artefacts will be reviewed, to improve fluency and familiarity with assessment tasks, thus increasing attainment. On the other hand, work by O'Donovan et al (2001; 2004; 2008) and Rust et al. (2003), focuses on developing students understanding of assessment criteria. They contend that if students understand how they are to be assessed this will improve their engagement with the assessment task and assist them in making improvements or adjustments to their assessment artefact, leading to better achievement. Shay (2004) identified that when students understood the assessment criteria they would have a better grasp of what would be valued in the assessment artefact.

This approach to supporting the development of student assessment literacy is an attempt to engage students as partners in the assessment process (QAA 2012). It then appears to be a legitimate exercise to try and assist students in understanding the assessment process from the practice of producing an assessment artefact to the approach that will be engaged to evaluating it. However as Sambell et al. (2012) suggest this entails much

more than simply showing the assessment rubric to students, but a commitment to devising ways of students learning how these rubrics would be applied in the evaluation of an assessment artefact. To bring the sentiments of the QAA (2013) to life, time needs to be made in the teaching for designing in an appreciation of assessment, application of assessment criteria and devising and interpreting of qualitative feedback. Developing assessment literacy in students is a complex process requiring integration into the whole learning journey of a unit, module and programme of study.

The notion of academic literacy can be equally applied to lecturers who need to become embedded in the ground rules concerning assessment in their department or subject of practice. Lecturers responsible for units of learning including defining module assessment are necessarily immersed in the requirements for demonstration of learning by the student (Jawitz 2009).

One approach to increasing assessment literacy of staff is to ensure that criteria are developed jointly involving lecturers who will be evaluating the assessment artefact (Saunders and Davis 1998). This approach would ensure that all lecturers start off with the same understanding. Further good practice suggested is to review the criteria regularly, to minimise the impact of potential misunderstandings or misinterpretations which can occur over time. Elander and Hardman (2002) replicate this process in preparation for an investigation of psychology examination marking. They developed criteria for assessment in conjunction with departmental staff, consisting of elements identified as important factors which could be evaluated in an assessment artefact. These criteria were used to increase staff assessment literacy and potentially reliability in marking and grading. These criteria were then used to promote student assessment literacy by providing explanations of how they should be used to ensure that evaluators can identify those components in their assessment artefact submissions.

Wimshurst & Manning (2012) apply a different model to develop student assessment literacy. Instead of providing students with examples or assessment criteria in isolation they developed an activity whereby students were able to review a range of exemplars of previous work submitted for their



same assessment task. They contend that this removes the anxiety of peer review of their colleagues which was disliked in the Bloxham and West (2007) study. The activity was where students had to read three examples, provide up to 150 words of qualitative feedback and their estimation of a mark or grade to be awarded. The thinking is that students will become immersed in the lived experience of defining of criteria and assessment requirements by reading and exploring previous submissions. Students improved their marks for their subsequent summative work, across the range of student ability, therefore presenting the utility of this approach to improve student assessment literacy.

## SECTION 2

### 1.2.1. RELEVANCE OF COMMUNITIES OF PRACTICE FOR ACHIEVING MARKING AND GRADING

There is a growing body of literature which explores the concept of communities of practice in higher education, some of which explicitly explores the induction and socialisation of new academics to their work context (Trowler and Knight 2000; Garrow and Tawse 2009). It seems that the original idea behind a community of practice was promulgated by Wenger (2000) and introduced a new understanding to how groups of people could work together with an increased level of performance and creativity that had a greater meaning than an organisationally constructed team. The group or community of individuals were described as such because of three specific aspects, that of *'mutual engagement, a sense of joint enterprise and a shared repertoire of communal resources including sources of learning and knowing'* (pg 229), all subject to the group of individuals performing together. The community would thus have a well-developed sense of purpose which was common amongst its members influencing their community identity. Whilst Wenger (1998) claimed that not every team with joint forms of working could be described as a 'community of practice', those who had a sense of place or belonging and socially accepted ways of resolving their differences which were unique to that group of individuals could potentially be characterised as such.

Gherardi (2009) has a description which could more easily be applied to lecturers entering into the academe, but on a more local perspective, marking and grading with a specific group of colleagues, this could be module, discipline or programme based. The term they use was that of a community of practitioners which lends more weight to the impact of the practitioners in shaping the practice rather than it simply being the group or community. Thus changing institution, although teaching the same discipline, will result in differing conventions amongst that community which must be learned and internalised in order to become a fully functional and accepted member.

Relevance of this concept to the activity of marking and grading is based around what Gascoigne and Thornton (2014) define as knowledge of the unspoken rules ('know how'), and practices which shape the application of the 'know that', and are unique to that community of practice. When considering how marking and grading of assessment artefacts is undertaken via either the use of normative referencing or the application of criterion referencing, the preceding discussion has identified some of the challenges. These include developing an understanding of the tools to be used, which is reliant upon the level of assessment literacy amongst staff to actualise the use of such tools, and the role of expertness in conducting assessment artefact evaluation.

Research has demonstrated that those new to evaluation of assessment artefacts are not necessarily less effective at applying set criteria or standards (Price 2005), this is in spite of being on the periphery of the module community of practice. One potential explanation was the seeming difficulty module leaders had in communicating the requirements of their assessment, which matched the inability to articulate the knowledge requirements by groups of markers. Price (2005) identified this as being illustrative of the difficult nature of tacit knowledge, remaining inexplicable but demonstrable in the tangible exemplar artefacts, of either model answers or previous submissions with marks and feedback.

Regular review of the criteria or rubric used for marking and grading needs to include all lecturers undertaking evaluation to facilitate discussion and debate about the content of the criteria in use (Saunders and Davis 1998). As well as improving consistency in application amongst lecturers, this provides an opportunity for those new to the community of practitioners to develop understanding of the implicit understandings of others in their team, contributing to the notion of a community of practice. Garrow and Tawse (2009) concluded that induction into local communities of practice for new academics was an important step to facilitate their engagement into the practise of their new community, moving them from a position of peripheral participation to full immersion and membership.

### 1.2.2. TACIT BELIEFS IN RELATION TO MARKING AND GRADING

Tacit knowledge or knowing has been described as:

*“ ... actual knowledge that is indeterminate, in the sense that it’s content cannot be explicitly stated.”*

(Polanyi 1969: p 4)

It is being able to carry out skills, tasks, roles and routines where individuals are only ‘subsidiarily’ aware of the process being undertaken and as such this is sufficient to prevent articulation of it. Polanyi (2009) also holds that tacit knowledge *“can be discovered without the individual being able to identify what they have come to know”*. The example often quoted is learning to ride a bike, once the skill is mastered, it seems impossible to articulate the mechanisms required to ride it, i.e. the how to, of how to ride. A standard approach to diminishing the quality of tacit knowledge, because it often cannot be gathered and arranged in a verbal form, is to place it in opposition to explicit knowledge. Explicit knowledge by contrast is then something that can be shared and as such takes on an objective and unbiased quality, able to be reviewed and ruminated upon by any observer. Tacit knowledge by contrast is personal rather than public knowledge, which cannot be verbalised, in so far as there has been no codification and it is context dependant, being demonstrated through practical application (Entwistle 2008; Gascoigne and Thornton 2014). The individual can express

the knowledge through practical application, which may remain indescribable, yet concrete results of its use can be seen.

Sadler (2009c) proposes in relation to marking and grading that not all measures used to undertake assessment artefact evaluation are communicable. They propose the existence of a tacit level of knowledge, which can come to be known by members of a group who have shared experiences, and is difficult to convey to those who are outside of this collective context. The exact mechanisms by which lecturers undertake the evaluation of an assessment artefact can only be revealed and communicated through observation and close contact by less experienced lecturers. Ajjawi and Higgs (2007) researched physiotherapists to try and reveal the methods by which they undertook clinical decision making and reasoning and how they were then able to communicate this to more junior colleagues. They identify how processes are subconscious but also context dependent, confirming Sadler's (2009c) proposition.

This illustrates the usefulness of exploring the concepts of 'know how' and 'knowing that', with the former helping to ensure that the latter is able to be put into action (Wyatt-Smith et al. 2010). Duguid (2005) maintains that increased quantities of 'know that' will not lead to 'know how', because 'know that' is developed through gathering explicit and encodable information, whereas 'know how' is only achievable through practice. The tacit component of knowledge is the factor which is unexplainable, indescribable and not found in a list of rules or set of guidelines. This explains why 'knowing how' cannot be translated and transmitted via facts or figures and not in a format which would be characteristic of delivery of such information. Polanyi (1966) goes so far as to suggest that within every act of knowing there is a tacit contribution.

The ability to transfer knowledge to new arrivals requires more than the transfer of codified knowledge. New arrivals need to gain a tacit understanding of the ground rules operated by their local Communities of Practice (CoP) to enable them to interpret information provided to them (Trowler and Knight 2000; Garrow and Tawse 2009). This is seen as '*a facet*

*of interpretation which originates outside the text to be interpreted'* (Duguid 2005). Once the newcomer is inculcated with the tacit understanding of modes of operation within their CoP, interpretation of the text will take place within these local norms. Without this inculcation the text or artefact could be approached from diametrically opposed stances being viewed as either truthful or disingenuous, with content of the text or source of information having no control of how it is interpreted.

One issue that has been explored is the continued variation in evaluation of assessment artefacts (Bettany-Saltikov et al. 2009; Hunter and Docherty 2009; Bloxham et al. 2015). Read et al. (2005) explored this from the context of examining gender differences in assessment. Their research reported on their findings from evaluations, carried out by 50 lecturers at 24 different HE institutions, of two history essays. Their findings suggested the reliability of marking and grading cannot be assured as the variation in achievement for these same two essays was between a high 2:1 and a fail grade. It seems that the estimation of quality of the writing is not achieved by an objective assessment, but rather this is mediated by the lecturer's notion or understanding of quality. Each individual lecturer's personal knowledge on what they understand quality to be, via tacit knowing has an impact on the evaluation of the essay. This is consistent with Leathwood (2005) who identified that different lecturers held different interpretations effecting the reliability of assessment outcomes. This situation is exacerbated by increases in the variety of approaches for assessment, each requiring a different approach for evaluation with individual lecturer's interpretations being applied, stemming from personal tacit knowledge and impacting on final evaluation.

The relevance of tacit knowledge to marking or grading assessment artefacts is how personal knowledge or internalised classification systems (Butler Shay 2004; Hudson et al. 2015) can impact upon the different interpretations and evaluations of that artefact. Interpretation is also dependent upon the rules, guidelines, ways of knowing and acting of their CoP, this results in the 'know how' being a product of that local CoP. Garrow (2009) concludes that

induction into local communities of practice for new academics is an important step. Even with this induction, evaluation of assessment artefacts will necessarily achieve different results, in the same and in different departments or institutions (Price 2005; Read et al. 2005; Bloxham et al. 2011).

### 1.2.3. EXPERTNESS AND EXPERT PRACTICE

Expertise is exhibited by an individual who possesses a great deal of knowledge and skill in relation to a discrete area or speciality (Goldman 2016). It is generally accepted that experts possess knowledge which is in excess of that expected by a lay person or 'novice' recently introduced to an area of practice. The concept of expert practice or expertise exists in most professional occupations. Take for instance medical practitioners in the realm of healthcare, where a consultant is seen as the expert (Ericsson 2004). In the area of accidents, an accident investigator and loss adjuster would also be viewed as an expert. Whilst in relation to law it is not unusual for there to be a reliance upon Expert Witnesses to present evidence in support of medical and other negligence claims.

Lecturers with expert subject knowledge are able to conduct the evaluation of the content of an assessment artefact by virtue of that extended knowledge in a particular field. Such in depth subject knowledge will not be the domain of every lecturer for the topic areas which they become responsible for evaluating. Alongside this, despite exemplifying increased subject knowledge, not every subject expert will be able to conduct effective artefact evaluations, this is by virtue of possessing limited expertise in the marking or grading process. This point highlights that there are a different set of skills to become proficient at which lie outside discipline expertise and related to the 'know of how' of the theoretical, procedural and practical aspects of marking and grading (Yorke et al. 2000).

Research by Saunders and Davies (1998) reported on differences in artefact evaluation which appeared to be related to the relative experience of lecturers. This perception was linked to two distinct areas of practice, the length of time taken to complete an evaluation of the artefact, and the subject

knowledge of the lecturer completing the marking and grading. Findings demonstrated that the longer it takes to evaluate the artefact, the lower the grade awarded, which was related to the experience of completing the marking and grading of this particular type of assessment. In essence, less experience resulted in a slower and lower evaluation of the assessment artefact. This could be a factor in the variable marks and grades which are commonly identified through a process of second marking or second consideration. Sadler (2009c) considers that the process of judgement used for assessment of an artefact, cannot be reduced to a set of identifiable '*measures or formal procedures*' (p160). If they existed they could then be applied by a lecturer with limited experience or exposure to marking and grading events, facilitating arrival at a correct evaluation of an assessment artefact. It is speculated that experts display a level of connoisseurship, with ability to undertake a holistic qualitative appraisal and include a valid and reasonable explanation of their judgement without recourse to set criteria (Sadler 2009d). This is presented as evaluators using pattern recognition, being able to quickly compare elements in the artefact with the requirements of the assessment task or published criteria (Brooks 2012). It is then easy to see why employing expertise in relation to marking and grading processes would be a useful concept to explore.

Saunders and Davies (1998) underlines differences between subject and non-subject experts and highlights a variation in being able to undertake an accurate evaluation of subject specific literature presented in the artefact. Limited topic knowledge resulted in lecturers focusing on more generic areas to facilitate the evaluation of the assessment artefact. As Sadler (2009c) suggests there are a variety of complex judgements that need to be undertaken across a heterogeneous sample of assessment artefacts, adding to the difficulty in completing the task. Substantial differences in artefact evaluation between subject and non-subject experts are predicable with students potentially fairing significantly worse under a subject expert with increased knowledge and expectations of what is presented. To combat this Saunders and Davies (1998) proposes that there should be workshops where the application of criteria are discussed, this should provide a forum

for those with experience to be reminded of the standard required, with potential for a reduction in a rigid approach to the application of their internal standards predicated by their expert subject knowledge.

How practitioners develop the level of skill and experience that would afford being awarded the title of expert has proved to be a perplexing concept. Several authors have developed a model of skill acquisition which they contend can describe how individuals progress through a number of stages of development. Table 2 explores three such authors' (Dreyfus and Dreyfus 1980; Hoffman 1998; Dreyfus 2004) approaches to explaining how skills are developed, presenting a theoretical perspective on skill development and show transition between these levels. Each of the authors have varying terms for moving from one stage or level to the other. The commonality being for each author there are only five stages to be negotiated, recognising there are no specific time limits attached to remaining at each stage or level. It is acknowledged that it is unlikely that all will reach the level of Mastery, Master and Expertise as defined by Dreyfus and Dreyfus (1980) and others.

### Comparison of Skill Acquisition Models

Stages of skill acquisition	<u>Dreyfus and Dreyfus 1980, Hoffman 1998, Dreyfus 2004</u> Description of similarities in stages of skill acquisition (differences per author are highlighted)
<u>Novice, Novice, Novice</u>	Includes individuals who are new, unfamiliar with the sphere of practice. <b>Ignorant of the domain. Require induction, including explicit sets of rules for action.</b> <u>Actions monitored and regulated by feedback - self and external evaluation, conforming to local Community of Practice (CoP)</u>
<u>Competence, Apprentice, Advanced Beginner</u>	Immersion in the domain learning through contact with real situations, supported by a mentor who highlights patterns, which after a time are recognisable and no longer abstract, are <u>contextualised and become meaningful</u> . <b>Apprenticeship lasts 1 - 12 years.</b>



<p><u>Proficiency</u>, <b>Journeyman</b>, <i>Competence</i></p>	<p>After increased experience can function effectively, working within rules and guidelines. <i>Individuals learn to identify the important elements which guide decision making. Performance is unsupervised, workers are reliable with a high level of functioning. May never progress beyond this level. <u>Decision making is context driven</u>, rules or guidelines are found to be lacking for more complex situations, decision making is then an individual act.</i></p>
<p><u>Expertise</u>, <b>Expert</b>, <i>Proficiency</i></p>	<p>Skills and knowledge derived from extensive experience in the domains. <u>No longer rule based decision making</u>, instead based upon recognition, <u>possess an intuitive</u>. <b>Demonstrates greater economy in the conscious effort required to complete the task</b>; <i>Increased emotional connections leading to anxiety as greater involvement in decision making informed by ‘situational discriminations’, choices relevant to the context.</i></p>
<p><u>Mastery</u>, <b>Master</b>, <i>Expertise</i></p>	<p><u>No conscious effort required, demonstrates instantaneous decision making, without recourse to rules or guidelines, reacts to signals or cues eliciting an appropriate response</u>. <b>A member of an elite group of experts who set the rules, standards, regulations. These individuals regarded as experts by other experts as well as their colleagues or peer group. Seen as qualified to teach newcomers or those at a lower level.</b> <i>Possesses the ability to make decisions which demonstrate intelligent and sophisticated discrimination, with an intuitive response which differentiates these individuals from the expert or proficient performer.</i></p>

**Table 2 Comparison of Skill Acquisition Models**

Benner (1984) author of a seminal text for nurses ‘From Novice to Expert’ adapted the skills acquisition model from Dreyfus and Dreyfus (1980) to facilitate exploration of data retrieved through their interviews with nurses. They mapped their findings against this skills acquisition continuum. Benner’s work became much more well known in health care settings for

devising a structure by which the development of nurses and others could be measured. This was not the researcher's primary purpose, but has become distorted over time (Gardner 2012) to being solely remembered for the devising of the 'Novice to Expert' structure, when it was simply used as a framework for analysing the research data.

Other authors approach the development of skills and know how from a different perspective, one which sees learning as a consequence of a holistic endeavour. Through investigation, Lave and Wenger (1991) suggest a process by which new arrivals are integrated into existing communities of practice (CoP) in a work context. They propose a position of legitimate peripheral participation where new arrivals have not been integrated into, or possess a sense of belonging to the department they have joined. Fuller et al. (2005) explain how the period from legitimate peripheral participation to full membership of a CoP can take varied lengths of time with the outcome for the individual being seen as potentially commensurate with this. In some cases the period of apprenticeship (which describes legitimate peripheral participation) lasted from between 1 -4 years. In that study this is characterised by a period of slow and protected maturation. New arrivals can and do begin to contribute, either positively or negatively, to the material resources of the department or group working context as a whole. New arrivals need to learn about their new work context, becoming integrated with the processes of becoming a full member of the CoP, including assimilation of the ways and ideas of their sociocultural practices. This process incorporates the learning of knowledge and skills relevant to the execution of functions for that particular CoP. Lave and Wenger (1991) propose an explanation of how initiates are developed into productive members of CoP. They suggest a non-linear perspective in opposition to those devised by authors highlighted in Table 2. These authors consider that acquisition of necessary requirements occurs via simultaneous action rather than via step by step approach. They propose a horizontal structure, seeing the development of relevant knowledge, skills and sociocultural capital as occurring simultaneously rather than via a hierarchical mode offered by other authors.

Individuals identified as experts in any domain share one commonality, the ability to portray automaticity and routinization in the completion of familiar tasks (Berliner 2004). This notion of expertness is explored by Crisp (2010b) who agrees that in relation to marking and grading, experts have the capacity to make speedy pattern comparisons between the assessment artefact and pre held patterns of model responses. They are able to identify patterns within their working context, responding in ways which demonstrate flexible application of knowledge. Elander and Hardman (2002) explain this as experts being able to detect and respond to information and stimuli or cues in their environment. These authors suggest that experts are able to gather and manipulate the available cues into smaller malleable dimensions, facilitating speedier information processing. They are able to segregate important and new information via access to their short term memory, with the long term memory being used as an underpinning framework to support decision making. This explains how those with years of experience can make a quick and accurate evaluation based on access to knowledge of previous assessment artefacts and the schema of a good response. If this is what is presented for assessment, no deliberation is required (Crisp 2010b).

Such actions set these individuals apart from those new to the profession or work context, defined here as 'novices'. In the sphere of marking and grading, lecturers with greater experience of being immersed in the evaluation of assessment artefacts are able to utilise and apply a greater resource of richer and more in-depth information to the task. Berliner (2004) also highlights that this level of expertise is not necessarily transferred beyond their specific practice domain and context, within which hours of experience through practice has been gained. The situated learning proposed here identifies why experienced new comers will need to again become immersed in their new CoP to demonstrate the expertise illustrated in a previous context.

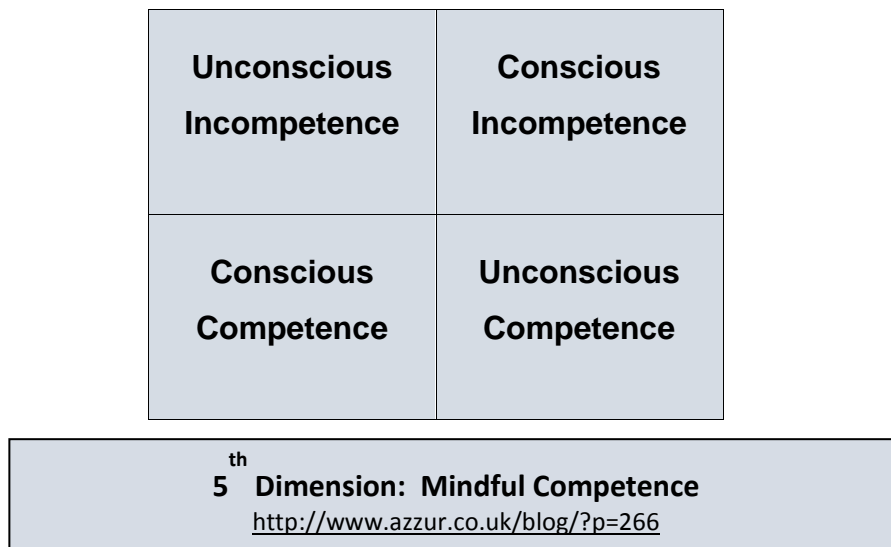
A further approach to describing skill and knowledge acquisition has been the competence quadrant. Whilst not attributable to a particular original author, Gordon Training International (2012), amongst others, have taken

credit for its production, it has proved a useful tool for visualising the journey from 'novice' to expert practice. The quadrant initially consisted of four phases: unconscious incompetence, conscious incompetence, conscious competence, with the final phase being unconscious competence (Figure 1). These phases can be mapped to those proposed by authors of skill acquisition models in the following ways:

- Unconscious incompetence - as a novice
- Conscious incompetence - as the apprentice or advanced beginner
- Conscious competence - as a state of competence or journeyman
- Unconscious competence - as a state of expertise or proficiency

However, a competence ladder has been proposed which includes a fifth step or dimension, adding to the initial competence quadrant. This 5th step is one where the expert achieves a process called mindful competence (Azzur 2010), where the individual is alert to the situation and to their approach to it. When mapped to the skill acquisition model, the addition by Azzur (2010) of mindful competence holds the position of mastery or becomes the master. In the context of marking and grading this would be evident in their approach to and execution of the task and in their ability to coach others in this skill development. A mindful approach is purposefully directing attention (Meiklejohn et al. 2012) to the task of marking and grading, noting thoughts and feelings, and then actions evoked by the assessment artefact. Mindful competence (Azzur 2010) goes beyond the master by paying attention to the cues which impact upon decision making about the artefact. This alertness to cues revealing how or why decisions are taken can be contrasted with the locus of a master or one who demonstrates mastery, which is not required as the role is completed through unconscious effort. Mastery as defined by Hoffman (1998) presents as practice of a skill without deliberate action, it is portrayed as an automatic function. As individuals learn more, the effort to undertake the skill reduces until what was once explicit becomes implicit and tacit.

**Figure 1 Competence quadrant**



Authors such as Gourlay (2011) and Northedge (2003) highlight the difficulties with adjusting to new roles and the need to develop competence in an unfamiliar area, as particularly challenging for practitioners entering academia. It is proposed that individuals being introduced to a new domain find difficulty in performing the requirements, but with support will develop and can engage in independent execution of tasks, albeit to an average level. Some individuals may develop at a faster rate and then continue to improve in both knowledge and competency or proficiency in the given domain, ascertaining expertise according to Ericsson (2004) and Ericsson et al. (2006). Demonstrating that the achievement of expertise is over and above being able to perform a task or skill in a competent manner or being able to show proficiency in it. The skill acquisition models (Table 2) explore each of the identified potential phases of development. When considering workload, lecturers reaching the step of being 'Proficient' (Dreyfus and Dreyfus 1980) or 'Journeyman' (Hoffman 1998) and 'Competence' (Dreyfus 2004) in relation to marking and grading would be the minimum requirement to become a fully participative member of their local CoP. To set this discussion into context, when lecturers are first introduced to marking and grading it is expected that they will find the task difficult to execute, completing it with varying degrees of accuracy.

## Summary

The literature review presented here was one crafted to focus on the areas which were thought to be of most pertinence to the chosen topic, exploring lecturers marking and grading practice in HE. As stated at the outset an iterative approach has been observed with a return to the literature to refine it, ensuring a connection with the final research findings. The importance of identifying what was already known about how lecturers completed the task of marking and grading, including what tools they used and the evidence base related to these was attended to. This identified limited understanding of the processes completed by lecturers to reach their final evaluations of student assessment submissions. This perspective therefore informed the research objectives of exploring marking and grading practices and highlighting any extraneous influences which may impact upon marking and grading practice.

Focus is then turned toward the utilisation of normative or criterion referencing for completing an evaluation of the assessment artefact. A case is made for and against both systems with authors being unable to agree. A possible antidote to the variations in evaluation of achievement of assessment artefacts, would be the adoption of a consistent approach, with the use of either system by all involved. Emphasis on assessment literacy has been largely focused on the student, however, this review has identified that staff assessment literacy is also an important aspect to be considered within the complexities of marking and grading. The final section of the literature review comprises of three separate but inter related sections. An exploration of communities of practice or communities of practitioners and the potential relevance of this theoretical concept to lecturers' marking and grading practice. This then leads to an unpicking of tacit knowledge and understanding, again ensuring that this theoretical concept is reviewed in relation to the marking and grading complex. The final section explores one of the components of this thesis title and that is the existence of expert practice and how such expertise can be gained. In the main, exploring this notion of development has been through the domain of skill acquisition, with the discussion returning to communities of practice through consideration of

idea of legitimate peripheral participation (Lave and Wenger 1991) being applied to newcomers or 'novices' of marking and grading.

The next Chapter (2) presents an exploration of the research philosophy which sets out my approach to the research and provides secure foundations from which the research takes shape and place. This focused literature review has served to ensure that my study is bounded enabling a deeper exploration of the issues relevant to lecturers marking and grading in HE.

# Chapter 2 PHILOSOPHY AND METHODOLOGY

This chapter explores two inter related aspects, my philosophical orientation along with methodological considerations, establishing their influence on my approach to selection of research method, the subsequent approach to data analysis and reporting of findings. The research methods and findings are explored in detail in the following Chapters 3 - 6. Finally, this chapter highlights the ethical considerations and decisions taken, including those raised by undertaking research within my own area of practice.

## PHILOSOPHICAL ORIENTATION

*"Today, if you are not confused, you are just not thinking clearly. In the confusion, there is tremendous opportunity, if only we are courageous enough to stay in it with eyes wide open."*

(Irene Peter)

The overarching purpose of the thesis is to present an investigation of marking and grading of Higher Education students' assessment artefacts by lecturing staff. The previous chapter has established the necessity for such an exploration and this section will explore more closely the philosophical underpinnings upon which this research is based. It outlines key assumptions that guided the decisions made, which ultimately shaped the final research product.

### 2.1 ONTOLOGY AND EPISTEMOLOGY

To investigate processes related to marking and grading in a fair and just way, there is a need to expound my ontological and epistemological orientation, exploring assumptions which are based upon this foundation. Ontology is about the very nature of reality and its characteristics (Creswell 2012), whilst epistemology is the study of knowledge, how we know what we know (Crotty 1998). Whilst these are subjects that I have previously given limited attention to, gaining a deeper insight into ontology and epistemology became a companion in the conduct of this research, ensuring that findings are credible and dependable (Koch and Harrington 1998).



Engaging with a philosophical approach facilitated me in questioning pre held assumptions about the subject of marking and grading in a way that I would otherwise not have done. In my role as a lecturer, marking and grading as a concept appeared unproblematic, it was only the actual practice which appeared to be more challenging.

Interviews were selected to explore the uniqueness of marking and grading practice from the perspective of lecturers, gaining insight into the features, disposition and reality of practise, exploiting a phenomenological method. Van Manen (1990) suggests that good phenomenological descriptions of experience can reveal their very essence allowing a window into an experience that has up until this point been concealed from view. Further exploration of this approach is presented in section 2.4. In the case of this research it permits the development of a new perspective on issues concerned with marking and grading.

Knowing then is derived from engagement with the purpose of study. Engagement with contemporary knowledge moves the researcher from a position of ownership to one of embodiment (Shotter 1993). Knowledge then is part of who the researcher is and of what they do. Implications of this lie in how findings of this phenomenological research study are analysed, reported and disseminated.

Cromby and Nightingale (1999) suggested that knowledge is so bound to an activity that it would emerge as a product of that activity. Knowledge of the practice of marking and grading therefore can best be derived from and through exploration of that practice. Investigation of this practice occurred through close connection and interaction with interviewees.

Interaction with the data as a practitioner (see Glossary p xiv) seemed inevitable; I am a researcher and a lecturer at one of the data collection sites. Interaction as a practitioner has facilitated the development of useful understanding and interpretation, along with the development of '*Phronesis*' or practical wisdom, as proposed by Aristotle (William 2008). This constitutes the exposing of knowledge which has the potential to be good or

beneficial when interpreted in light of a specific practice context. This highlights the importance of acknowledging a practitioner research element to this study. Orr and Bennett (2009) present a compelling argument supporting the advantages of practitioner or insider research. They report on issues of trust and the effect of unfamiliarity on the research process, with practitioners being able to build rapport with interviewees much more quickly, thereby improving the data retrieved. The practical application of this will be explored further when reviewing methodology and methods (Section 2.4 and Section 2.5). Practitioner research was used to reveal knowledge that would lead to practice development through the uncovering of practical wisdom, a position supported by Brooker and Macpherson (1999). This culminated in the development of a supportive mechanism to guide practitioners in marking and grading.

## 2.2 PRACTICE OF MARKING AND GRADING

The practice of marking and grading, the awarding of a mark or grade to an artefact during a process of evaluation could be viewed as a socially constructed activity. This activity is culturally and historically bound and it is one which occurs frequently in Higher Education and other education settings. This perspective is acknowledged by authors such as Bloxham and Boyd (2012), Shay (2004, 2005) and Rust et al. (2005). It is an activity in which assessors review work produced by students, giving it a value, estimating its worth, by providing a mark or grade. This activity may result in different values being awarded at different times, by a different marker or grader or indeed the same individual, thus seen as cultural and historical specificity (Pinot de Moira et al. 2002; Bettany-Saltikov et al. 2009). Reasons for this variation are multi-factorial, though could be a direct consequence of differing interpretations that are culturally and historically bound, within particular communities of practice. There has been an upward trend in school years education of requests for remarking of examination papers by Ofqual (Office of Qualification and Examination Regulation 2016). In 2015 there was an increase in such requests which was 27% greater than the 2014 level. Of the requests made for remarking 18.9% saw a grade

change (Ofqual 2016; Catlow & Fisher 2013). Subjectivity seems to be accepted as an inherent facet of the evaluation processes undertaken during marking and grading.

Outputs from the process of marking and grading are constructed by markers or graders who are reliant upon their individual way of understanding the world. Burr (2003) and Gibbs (2016) alludes to this notion, as it appears that there may be more than a single way of seeing or evaluating student work which is presented for assessment.

The preceding literature review (Chapter 8) demonstrates that processes connected with marking and grading have been of concern however there is limited evidence of an exploration from the perspective of lecturers completing the process. The literature also demonstrates limited evidence on the processes employed for developing novice markers including how they would then be inducted into their role of evaluating student assessment artefacts. This raises a concern regarding how novice and other evaluators are prepared. Indeed this raises questions about what is the best way to achieve this and thus eliciting if there are experts in the field from whom the practice can be learned. Attending to the components identified in the literature review has demonstrated a limited holistic understanding of the processes undertaken during marking and grading. This seemed then to be an area worthy of deeper exploration, identifying all of the elements completed by lecturers when marking or grading rather than simply focusing on specific issue for example application of criteria. The research questions have purposely remained broad enough to uncover what evaluators do rather than focusing on pre-determined areas that are already deemed to be important. This became a distinctive part of my approach to this enquiry.

## 2.3 AIMS, OBJECTIVES AND QUESTIONS

Aims, objectives and questions for the research arise from wanting to explore the social world of marking and grading, in such ways which are unmistakably representative of and related to activities we complete (Cromby and Nightingale 1999).

### 2.3.1 AIMS

To investigate the process of marking and grading from the perspective of the marker. To recognise good marking and grading practice deriving a model of 'expert' practice aiding the development and proficiency of novice markers.

### 2.3.2 RESEARCH OBJECTIVES

1. Examine assessment by exploring marking and grading practices.
2. Identify what are 'good' marking practices to develop a concept of 'expertise' in marking and grading.
3. Explore novice lecturers thoughts on marking and grading
4. Explore cognitive processes and extraneous influences on marking and grading practice.

### 2.3.3 RESEARCH QUESTIONS

1. What do lecturers believe they do when they are undertaking marking and grading?
2. What thoughts or cognitive processes occur in lecturers during the act of marking and grading?
3. What exists to ensure equity or consistency in professional or inter-professional marking and grading, this includes inter-rater reliability.

## **METHODOLOGICAL STANCE**

*“Progress lies not in enhancing what is, but in advancing toward what will be.”*

Kahilil Gilbran – Handful of Sand

This section explores the research methodology utilised when gathering data concerned with marking and grading. Close attention to the philosophical underpinnings, as already highlighted, facilitated a good match between methodology and method. Ethical considerations closely aligned to the study design are deliberated. The overall aim was to enable a clear link to be made between the aim of this study and resultant approach. There will

follow an explanation of the selected research methods for data collection which are presented directly prior to presentation of the findings. Therefore Chapters 3 and 4 will explore method and findings from Cognitive Interviews; Chapters 5 and 6 will explore method and findings from Protocol Analysis Interviews.

## 2.4 RESEARCH APPROACH

The research approach was guided by the preceding discussion on ontology and epistemology. The selected approaches were used to establish findings or knowledge of practical wisdom.

Whatever the research methodology, investigators themselves will influence the research process through their interactions, assumptions and experiences. Carpentier (2008) explains that whilst an illusion of neutrality and objectivity permeates quantitative research, quantitative data cannot exist as an entity alone. If this is taken to be the case, then quantitative data are also to be judged as are social constructs. Gilborn (2010) goes further by exploring reasons why quantitative research has maintained an aura of objectivity and the answer, he asserts, lies in the presentation of statistical conclusions. The use of numbers results in exclusion by creating a barrier that many would be research consumers do not feel able to understand or more importantly, not competent enough to challenge.

For the most part this research is positioned within the qualitative domain.

## 2.5 METHODOLOGICAL MEANS

Exploration of my ontological and epistemological assumptions has made clear a desire to uncover experiences of marking and grading from individuals undertaking that role. In the context of Higher Education this research is directed at exploring the experience of lecturers completing a common task, with inherent complexities. Such exploration was directed at illuminating the practice of marking and grading, uncovering potential issues

and identifying good practice. To achieve these aims consideration was given to the methodology upon which this investigation was constructed.

Methodology emanates from my espoused philosophical beliefs and thus pinpoints a scheme for conducting systematic enquiry. It facilitates contextualisation of the research. Thus it is the 'recipe' or 'roadmap' for completing the research in this thesis as it embraces philosophy, assumptions about validity, and explanation of preferred methods to be applied when collecting data. My philosophical perspective identifies that studying the lived experience of lecturers conducting marking and grading is of greatest interest. Exploring experiences from these individual actors, in their own words, will present a powerful account of what it is like to mark or grade student essays, which is a taken for granted activity in education. Such a position leads to utilisation of the philosophical tradition and research method of phenomenology (Van Manen 1990). Furthermore, the philosophical lens proffered directs toward Hermeneutic or Interpretive phenomenology.

### 2.5.1 PHENOMENOLOGY

Phenomenology was founded by Edmond Husserl, one of the 20<sup>th</sup> Century's most eminent philosophers (Sadala and Adorno 2002; Beyer 2013). His approach encompassed a belief that the nature of lived experience can only be uncovered through exchanges between the researcher and the researched. Use of communicative mediums of attentive listening, observation and interaction could result in the ability to construct more knowledgeable representations of reality (Wojnar and Swanson 2007). Husserl proposed a process of phenomenological reduction (Paley 1997) which is a total suspension in belief of anything in a world other than the phenomena to be examined.

In essence, Husserl requires abandonment of an ordinary outlook enabling the researcher to achieve a position of neutrality or '*Epoche*' (Beyer 2013), this results in researcher separation from the phenomena by setting all existing preconceived ideas of it aside. The researcher would successfully

achieve separation from the phenomena by metaphorically 'bracketing' their beliefs in a reality which is neither verified nor rejected (Koch 1995), an act of phenomenological reduction. Putting aside everyday assumptions, usual understanding and biases in the adopting of neutrality is claimed to be a way of guaranteeing that assertions made from study findings are representative of truth; bringing forth accurate depictions of the phenomena (Moustakas 1994). This approach does not reconcile with my perspective on practitioner research and the benefits for the research of utilising such a strategy (see 2.5.7).

'*Epoche*' does not fit with my ontological or epistemological position; hence a further branch of phenomenology was explored and then enlisted.

## 2.5.2 HERMENEUTIC PHENOMENOLOGY

This approach to phenomenology was proposed by Heidegger, Smith (2011) reports on Heidegger's departure from Husserl in relation to phenomenological thought. The significant difference originates from Heidegger's idea of '*being*' to create understanding of a phenomena (Dowling 2007). He proposes that researchers should approach their exploration, interpretation or appreciation of the phenomena through '*Dasein*' which means being there (Fleming et al. 2003). For Heidegger his interpretation of '*Dasein*' includes:

*"...an awareness of ones being and belonging to the world, availability and use of the world and relating to or with others."*

(Fleming et al. 2003: pg 117)

This translation clearly expresses consideration of the phenomenon incorporating the historical, cultural and social influences. A consideration of experiences of '*being*' as historically situated is viewed by Koch (1995) as a valid approach to contextualising knowledge and understanding gained to describe, interpret and perceive experiences of the phenomenon. This fits with how I positioned myself as a practitioner researcher. Pre-existing knowledge of the social, historical and contemporary context of marking and

grading was enlisted to gather and analyse the data, and then report on the research findings.

The goal of employing hermeneutic phenomenology in this study is to arrive at the '*Essence*' of the experience of marking and grading. This approach presented an opportunity to reveal tacit knowledge, that is knowledge that lecturers have difficulty in articulating in the experience of marking and grading (Ajjawi and Higgs 2007).

The chosen research methodology facilitated in-depth investigation enabling identification of lecturer's specific beliefs, values and traditions in relation to marking and grading of student work. Van Manen (1990) identifies hermeneutic phenomenology as a scientific study of persons. Contributors to research in this phenomenological tradition are not usually named subjects or participants. In this study contributors are named as lecturers and/or interviewees. Data collection was directed towards exploring their perspectives of marking and grading, noting their accounts and individual or unique insights. This is an important point, lecturers or interviewees were viewed as co-constructors of the new knowledge that was derived from the findings of this research venture. Insights gained may not be entirely comparable, and as such may be viewed as unique, limiting their scope for application beyond the research context. Important insights into the lived experiences of markers and graders (lecturers) were gained through a hermeneutic approach to collecting data, individual interviews were selected as the vehicle of choice to facilitate this.

### 2.5.3 ACTIVE INTERVIEWING

Holstein and Gubrium's (1995) process of Active Interviewing was adopted as it provides a forum for construction of detailed descriptions in the form of storytelling. Interviewing and the interview process is not viewed as one where the researcher has to unearth what lies hidden within the research subject. It is simply seen as an opportunity for narrative exchange where the researcher asks questions which prompt the interviewee to reflect on their experiences in response to specific objectives, framed by the research



questions. I expected the interviewee to respond, relaying their story which was contextualised to a particular time, space and event. The interviewee is not inactive in this process; they are able to pursue their individual interpretation of the demands of the interviewer. The interviewee can continually weigh up the possibilities available to them, to best construct meaning from sections of experience which they are being called upon to recount. In the case of my research this was related to a marking or grading event, within a contemporary time frame, whilst responding in an interactional way to the immediate circumstance of the interview. A sharing conversation ensued between myself and the interviewees, with responses being formulated in answer to particular questions, prompts, comments and requests for illumination. In this way the interviewee is directed into revealing detailed and particular aspects or themes of the experience of marking and grading. The interview process thereby becomes a reciprocal exchange where there was a dynamic interaction which culminated in the production of a meaningful narration or storytelling of a particular marking and grading experience. An individualised approach to question presentation is adopted and adjusted to suit the context of the interview, responsive to what the interviewer sees as requiring further illumination. In this approach neither the interviewer nor the interviewee adopted traditional researcher and subject roles (Rubin & Rubin 2005), of the researcher taking charge of the research interview, or the interviewee answering questions to complete the 'blank spaces' on the interview question sheet.

Using hermeneutic phenomenology facilitated the hearing of interviewees' stories, exploring their truths. To assist this exploration two methods of collecting data were utilised, these methods are explored in more detail in Chapters 3 and 5. The rationale for using these two methods was to ensure that this research was able to elicit a breadth and depth of information which would not be possible had one of these methods alone been used. When reviewing Chapters 3 and 5 the strengths of one method are indeed seen as a weakness of the other and vice versa. Hence the strength of my research is the utilisation of two methods with a variety of interviewees across all of the research sites and among all levels of lecturer experience. Marking and

grading is commonly learned through a heuristic approach, this study explored experiences, uncovering a range of features of marking and grading through the story telling of interviewees.

#### 2.5.4 RESEARCHER REFLEXIVITY

Researcher reflexivity comes into play here, as Mason (2002a) highlights it consists of exploring the intersection of the role of the researcher and their centrality and “active construction in the collection, selection and interpretation of data” (Finlay and Gough 2008: ; pg 5). Indeed Koch (1995) and Gadamer, as referred to by Annells (1996), characterize reflexivity as an iterative process when researcher understanding and interpretation occur synergistically. Understanding cannot exist without interpretation.

Understanding comes from some pre-existing knowledge of the phenomena such that other researchers would undoubtedly ‘unfold a different story’ (Finlay and Gough 2008 p 5). I suggest, as do others (Van Manen 1990; Finlay 2002; Finlay and Gough 2008; Orr and Bennett 2009) that research is not only historically and culturally situated but is personally affected by the contributors. The researcher, research participants and their previous individual experiences of the phenomena under investigation, in this case the experience of marking and grading, are related in a complex interplay in co-constructing the research findings.

Engaging in reflexive practice or the use of reflexivity is the process which enables the researcher to usefully, critically and judiciously (Finlay 2006; Finlay and Gough 2008) incorporate researcher knowledge and understandings with the gathered data from research participants. This was done in such a way as to ensure that it is evident to the reader where ideas emanated from, as the researcher responded to the environment and situation of the enquiry.

Meaning can only be derived from what *‘is’*. The activity of marking and grading exists, albeit as a social construction. Through using the selected data collection methods it was contended that this would allow the meaning of this activity to be brought into view. Hermeneutic phenomenology was

thus an appropriate approach to explore what *'is'* in this context with researcher reflexivity being used as a part of a toolbox for data interpretation, and consequently construction of findings.

### 2.5.5 METHODOLOGY VS. METHOD

Chapters 3 and 5 will explore the methods of data collection employed in this research study. However some consideration here is also given over to the relationship between methodology and method. A case has been made thus far that the primary concern for this research was the exploration of lived experiences of lecturers undertaking marking and grading. It was therefore reasonable to propose and apply a method or methods which facilitated such investigation. One proposed method was to undertake Protocol Analysis (PA) interviews, which are conducted while an interviewee performs an authentic task, in the case of my research this entailed lecturers marking and grading a piece of written text. Capturing this process in action was an appropriate way to illuminate experiences of markers. Merleau-Ponty's (Flynn 2011) assertion about 'embodiment' was accepted, making embodiment in human action of great relevance. He argued that the 'body knows how' to do things, in this case the 'body knew how' to mark or grade student work for assessment. It was proposed that utilisation of a PA interview technique would capture embodiment in the form of marking and grading, facilitating an insight into what *'is'*.

It is accepted that incorporating innovative approaches in the construction of research findings may result in foregoing of some elements which would not need to be tolerated with more traditional methods. The use of Protocol Analysis (PA) (Chapter 5) has the potential to interrupt the process of marking and grading in an effort to gain access to knowledge which is usually inaccessible. The ability of the lecturer to perform the task of marking and grading could be impaired by the presence of recording equipment and indeed the researcher. Marking and grading of student work for assessment is normally undertaken as a solitary exercise and in almost total silence. Utilising PA may have been harmful to the authentic task of marking and grading, as it occurred with some alterations to usual lecturer

practice (considered further 2.6.5). However, it is asserted that this technique would allow contact with as near to embodied knowledge on marking and grading as could be, presently, possible. Merleau-Ponty is reported to view the accessing of this embodied knowledge as a means of reclaiming wisdom of the world being explored (Benner 2000; Flynn 2011; Thomas 2005).

A second data collection method, Cognitive Interviewing, was utilised. The technique used was much more researcher led than the Active Interview approach of Holstein and Gubruim (1995) described above and that of traditional qualitative research interviews with their unstructured question schedules. For CI interviewees participated in a process of deep post task reflection, returning the interviewee to the thoughts and feelings of the experience or in this research the marking and grading task. The technique draws on data gathered from the memory of the interviewee and as such from a social constructionist perspective meaning is constructed through engagement with the emerging data. A process of Active Interviewing acknowledges the interviewees central role in constructing the narrative around their engagement in the task being investigated. A two-step process consisting of iterative construction of interview data during questioning, involving the researcher interpreting and clarifying meaning; reconstruction occurred as a second process during the writing up and reporting phase. Van Manen (1990) sees the process of writing as an essential component of phenomenological investigation.

### 2.5.6 INTERVIEW SAMPLE

The determination of the number of interviewees and the method of selection is not fixed in the research literature. Therefore it seemed acceptable to utilise convenience sampling to select the cohort to contribute to the gathering of data. It is known that convenience sampling is not one of the strongest of techniques for establishing a research sample (Parahoo 2006) and as such it is accepted that my research cohort is not representative of the entire eligible population. The final cohort were interested in the topic and answered email advertisements, at their institution, which were used to

publicise the research. To ensure that I was able to maintain an overview of the composition of the cohort, a sampling frame (Appendix 3) was maintained. This provided detail of what type of interview was to be completed (either Protocol Analysis or Cognitive Interviewing) and whether the interviewee was a 'Novice' or 'Experienced' marker. Details of interviewees are provided early in the Findings Chapters (4 and 6) Interviewee Key 1 (4.2); Interviewee Key 2 (6.2).

### 2.5.7 PRACTITIONER OR INSIDER RESEARCH

When exploring potential research methods to be adopted, a practitioner research strand appeared to be key, as a focus for the research aim was the production of findings with some everyday applicability to practice (Brooker and Macpherson 1999). Practitioner research seeks to explore issues which are of direct relevance to practitioners (Anderson and Herr 1999). Thus what may be viewed as a separate strand of data collection, is in essence, something which was inextricably intertwined as an embodied feature (Thomas 2005) of being a practitioner researcher. There was no direct intention to formally collect further strands of data; however, informal discussions did occur with colleagues. These were focused on approaches to everyday marking and grading practice, including addressing practical concerns which arose during and in between the data collection and writing up phases. By the definition used here, this information does not constitute data. As the principle investigator and an insider or practitioner researcher in this study, it did not preclude me from participating in such conversations, which undoubtedly influenced the continued perception and interpretation of findings within the research study. It is contended that data, retrieved through Protocol Analysis interviews (PA) and Cognitive Interviewing interviews (CI), were interpreted through the researcher's contextual lens. This contextual lens or perspective was inevitably informed by connections with other lecturers engaged in marking and grading practice.

## 2.5.8 RESEARCH SITE AND POPULATION

A purposeful homogenous sample of lecturer's based in a Faculty of Health, or similar, at any one of the 4 participating institutions were recruited for interview. The sample of participating institutions consisted of two universities which pre dated the 1988 Education Reform Act and two who were post this date and had therefore benefited from the changes to Higher Education introduced by this Act (legislation.gov.uk).

A heterogeneous sample of lecturers from professionalised disciplines in health were recruited. They traditionally engage in a particular discourse relating to patient or client care (Clouder 2005) or safety (Armitage et al. 2011) promoting the notion of safe practice. The lecturers were from the following health professional backgrounds: nursing including adult, mental health, district nursing and palliative care, radiography, physiotherapy, paramedic practice. At the end of recruitment and data collection phases of the research, I considered that a sufficiently heterogeneous sample, in terms of experience, had been gathered, and given interviews.

## 2.6 ETHICAL CONSIDERATIONS

The ethical issues which are considered here relate specifically to executing the research methods described in Chapters 3 and 5. These include identification of research population, recruitment and sampling, gaining informed consent, assuring and maintaining confidentiality and safety of research data.

### 2.6.1 RESEARCH ETHICS

In considering the issues highlighted above, the emergence of ethical anxieties regarding collection and analysis of data became appreciable. Selected issues are examined here by the application of the 4 major ethical principles: autonomy, concerned with self-rule and independence in decision-making; beneficence, referring to kindness and goodness; non-maleficence, denoting to do no harm, act with no malice; justice, incorporating honesty, fairness, integrity and equality (Gillon 1994). Whilst these principles are

used by Gillon (1994) in the context of medical ethics, these 4 tenets are equally applicable to the conduct of research. An alignment between the research approach to ethics and the ethical principles which are familiar to the health professionals participating in this study is hence achieved. A number of further considerations are highlighted in relation to how these ethical principles came to be operationalized.

Virtue philosophy was engaged to ensure that ethical issues were given due consideration at each point in the study, importantly enabling an appreciation of moral or ethical values that were relevant within the given context.

Adherence to a rigid set of rules without careful consideration of the situation would result in ethical decisions which were not flexible enough to respond to the specific context (Soobrayan 2003; Kvale 2008). This approach enabled careful deliberation of my duty as a researcher to evaluate the consequences of my actions, ensuring an appreciation of my ethical responsibilities at each stage of the research process.

### 2.6.2 LECTURERS/INTERVIEWEES

Data collection for this study was achieved by close connection to lecturers via interview in my own and other Higher Education institutions. Whilst other authors have different terms to refer to those from which data are gathered (person or persons by Van Manen (1990) or co-researchers (Moustakas 1994)), for ease of reading to be consistent with my own community of practice lecturers participants in this study will be referred to as interviewees. A partnership between the researcher and researched is signified by referral to data collection rather than data generation; as insights into the lived experiences of lecturers marking and grading.

### 2.6.3 GAINING CONSENT

Individual informed consent was attained from each lecturer prior to the commencement of interviews. Informed consent assured both parties that each understood what was required from the other. This demonstrated respect for the first ethical principle of autonomy, a regard for self-rule. An information leaflet was provided to potential interviewees. This highlighted

the study purpose and was supported by verbal repetition of study requirements, immediately prior to data collection. This is comparable to a second stage in the consent processes, with the first stage being seen as acquiring University Research Ethics Committee approval. Gaining approval is receiving informed consent to undertake the study from the Institution (Ethics approval APPENDIX 1)

Detail of the procedure followed in recruitment is outlined in the Recruitment Protocol (APPENDIX 4). After initial email contact, an opportunity for verbal contact was arranged between the researcher and potential interviewee. Verbal contact provided an opening for clarification of details regarding the research approach and their expected contribution, which may still have been a cause for concern. The depth and breadth of information given revealed any risks and also potential benefits of contributing to the research (Kvale 2008).

A completed and signed consent form was acquired directly prior to commencement of the research interview, this detailed all their rights including their right to decline to participate. The right to withdraw from the study included the right to withhold data from being analysed and reported (Moustakas 1994) . If following data collection there was a decision to withdraw, interviewees would have been given two options with regard to their preferences for use of their data. As a researcher the most desirable option would be for information freely given to be used within the data analysis, including reporting of findings. The second option would only allow for information to be destroyed and therefore not used when reporting findings. Potential interviewees were assured that they were not obliged to provide a reason for withdrawing from the study, at any point in time. No interviewees decided to withdraw from the research study, therefore all data reported constitutes that of the recruited research cohort.

#### **2.6.4 POPULATION, RECRUITMENT, SAMPLING**

The identified population for this study are United Kingdom (UK) Higher Education lecturers. All lecturers who were teaching and completing marking



and grading in the primary language of communication, English, were eligible for recruitment. No provision was made in the calculation of research costs for interviews to be attended by a translator. This approach does not pose any threat to respect for the ethical principle of justice. Communication in English, in verbal and written formats is a pre-requisite for employment in Higher Education in the UK as is evident by the way new roles are advertised (Jobs.ac.uk 2013).

A respect for the autonomy of lecturers agreeing to participate was maintained and this included preserving confidentiality. When an interviewee belongs to a defined research population they will be over exposed to research issues, which are judged to be of concern to members of that population. Interviewees thus display an increased knowledge and awareness compared to their colleagues, which could pose a threat to their anonymity. Within the context of this research the defined population were lecturers in Higher Education, they are the only members of that environment who could be exploited for their particular knowledge regarding marking and grading. The literature review (Chapter 1) highlights research on different aspects of the processes and practice or act of marking and grading. There remains utility for this research to contribute to an extension of knowledge of this important task.

To gain access to a sample from the eligible population, applications were made to the relevant University Research Ethics panels for approval. Following this it became necessary to obtain further 'local' permission, to facilitate access to the potential research population. When attempting to recruit study participants at unfamiliar external institutions, a local conduit was required. The initial study proposal was to use Associate Dean for Learning & Teaching or similar to introduce the researcher to the potential population in the identified School or Faculty, from which the research sample, was to be drawn. The Associate Dean (de Witt and Ploeg 2006) was simply used as an intermediary, verifying the legitimacy of the research. Attention must be paid here to concepts of power and a potential perceived compulsion to participate as being encouraged by Associate Dean, who

when compared to the lecturer, may occupy an elevated hierarchical position.

In practice, after making contact with external institutions via the department or individual charged with organising ethical approval, I was then supplied with the name and contact details of a Head of Department (HoD) who then kindly distributed my email request across the Health Faculty. One HoD did distribute my email request to a particular cohort of new lecturers rather than the whole Faculty. Some recruitment then took place via the 'Snowballing' technique; occurring when one lecturer in an institution decided to participate, who then passed information on to their better known colleagues or friends. This was used as an auxiliary method to enhance the research sample (Noy 2008). This did pose a threat to internal confidentiality, over which the researcher had no control. External confidentiality (outside the department/institution) of the interviewee was maintained, with their identity being only by known the researcher.

An additional consideration was whether having pre-existing knowledge of the researcher would impact upon the voluntary element of participation. A decision to undertake research in my own institution and faculty could be viewed as a pragmatic decision. It did present particular ethical challenges. Where research participants were known to me there may have been some researcher effect which could have been both positive and negative (Mercer 2007, Anderson and Herr 1999). From an ethical perspective a pre-existing relationship between myself and the interviewee could have resulted in a feeling of greater compulsion to participate, along with feeling pressurised to give what they perceived to be the 'right' answers. A continuing respect for interviewee autonomy was emphasized by ensuring that at no time was anyone harassed into participating, all interviewees volunteered to take part out of an interest in the research topic. This is where my insider status, particularly at my own institution may have exerted particular impact. Interviewees' autonomy may be threatened by assumptions of my level of knowledge of the topic area and a perception of their own assessment and marking and grading literacy. A threat to their autonomy because of my

insider status cannot be avoided, since there was a long history of being associated with assessment; care was taken to respect the principles of beneficence and non-maleficence. This manifested itself in being sensitive to the interviewees level of understanding, by taking care not to belittle their approach to marking and grading, simply asking questions to ascertain what individuals did and why. This approach is commended by interviewees (Chapter 8.2). Guarding against prior personal connections impacting overtly upon data gathered and analysed was achieved by comparing data gathered at institutions where I was an unknown entity.

### 2.6.5 DATA COLLECTION

There was no wish to cause physical or psychological harm to research participants; though it was acknowledged that lecturers may find participating in this study a stressful experience. A respect for non-maleficence is part of the attitude and approach adopted. Care was taken to reassure interviewees of a focus on maintaining confidentiality. Interviews were undertaken at a pace that enabled the interviewee to understand the process and which they were comfortable with. During the PA interviews I remained present during the marking and grading process. It was acknowledged that an interviewee could have found this situation stressful, increasing their sense of psychological danger, intensifying any feelings of anxiety. Interviewees with a predisposition to nervousness might find this technique overly stressful and therefore would have chosen not to participate. At no time during any of the interviews did any interviewee express outward signs of anxiety, at the end of the interview a number of interviewees commented upon the ease with which the process occurred because of my attitude and approach taken. It is inescapable to think that having another person present, during Protocol Analysis, will have limited or no effect. When considering non-maleficence an increase in interviewee anxiety could occur, with them questioning their efficacy in completing the task. Interviewing and post protocol analysis debriefing encourage a deep level of introspection on a process which may be a well-established practice for that individual. Self-examination and reflection may lead interviewees to have doubts about their

individual assessment practices. Allowing time for interviewees to debrief and discuss their experience assisted in reducing any such anxiety.

As already alluded to, respect for autonomy of the interviewee was paramount, providing assurance that the process could be halted at any time, if necessary. Recruitment of lecturers to participate in the data collection did not exclude those lecturers with learning disabilities, mental health problems or a terminal illness, as the presence of these conditions is not uncommon in the general population and often remain hidden. It was at the discretion of the lecturer as to whether the existence of such conditions was disclosed to me. Whilst having appropriate concern for interviewee autonomy, I also had a duty to their well-being and as such maintained an acute concern for the principle of non-maleficence. If any undue or increased stress or anxiety had been detected during data collection, the interview process would have been terminated, ensuring appropriate support was gained for the individual. During the interviews to be reported upon in Chapters 3 – 6, no interviews were terminated by either party.

Interviewees agreeing to participate in the research by signing a consent form, were made aware, in principle, of the extra time commitment required to complete either data collection approach. This increase in workload is unavoidable, even when maintaining a respect for beneficence and non-maleficence. An appraisal of the ethical dimensions of the situation with consideration from a prima facie perspective ensured respect for interviewees' autonomy, above any other competing ethical principle. The interview would have been terminated if the workload had become too onerous.

Although the study does not have any direct involvement with the student population, work that they submit for assessment will play a significant role in identifying the nuances of marking and grading undertaken by interviewees. Using authentic student papers for summative assessment maintains the authenticity of the assessment process, it is important that lecturers know they are providing a mark or grade for a summatively assessed piece of work. This is where a noticeable shift toward a deontological perspective

informs the data collection process. A respect for the individual situation of each interviewee occurs in ensuring they are engaging in an authentic task, without deception or increase in workload for the individual or their colleagues. However, it was acknowledged that essay papers used during execution of Protocol Analysis interview technique may be affected by the data collection method. To minimise any disadvantage to students and to respect the ethical principle of justice, a stringent process of moderation of these papers will be put into place.

#### 2.6.6 CONFIDENTIALITY

Assurance of confidentiality for research participants is essential when maintaining respect for the ethical principle of beneficence (Orb et al. 2001). Processes with the potential to protect the identity of interviewees were put into place. For example the approach of employing gender non-specific pseudonyms, concealing both the name and gender of interviewees. This seemed to be an appropriate response to increasing the security around interviewee identity. When working and researching a small environment it becomes even more difficult to maintain confidentiality, with contributors being more easily identifiable. An emerging consideration with the increased use of technology is the presence of a digital footprint (Weaver and Gahegan 2007), where personal information about interviewees will be stored digitally. In this research digital footprints are created through reciprocal electronic communication, recording of information prior to interview, following interview via interview transcriptions and any analysis that I may conduct. The need for security and the right to privacy of individuals becomes ever more important in the face of technology (Weaver and Gahegan 2007).

Out of respect for beneficence and non-maleficence, all digitally stored files were password protected, stored in a password protected folder, on a server which invariably requires a password for access. Careful data storage included separation of data from lists of interviewees.

Devising composite stories was a strategy which was exploited to maintain the anonymity of transcription data. Responses and comments from a number of interviewees are joined together to present findings (Creswell 2012), a useful approach when constructing reports of findings for dissemination locally, nationally and internationally. This recognises and respects the ethical principle of justice which charges the researcher with acting in a fair and honest way, maintaining integrity. Presentation of interviewee responses in full would make them vulnerable to disclosure of their identity. Utilising the approach of composite stories, respects the ethical principles of beneficence and non-maleficence by protecting their anonymity and maintaining confidentiality (Hewitt 2007).

It is possible that I may become aware of poor or substandard marking and grading practices. Interviewees could be subjected to individual harm, if superiors or those with responsibility for appraisal became aware of practices which could be viewed as sub-standard. A simultaneous lack of respect for beneficence and non-maleficence would occur if I were to breach confidentiality of the interviewee and report this finding to individual managers or supervisors. It is important to maintain interviewees' anonymity by respecting their autonomy. This is a significant safeguard against a violation of interviewees' trust in the researcher, and research approach. The creation of composite stories enables the telling or reporting of interviewees' lived experiences without breaching confidentiality. A further issue is highlighted in that the risk of disclosure by the interviewee of something which could pose a significant threat to student success was not initially explored (Hewitt 2007). The setting of ground rules incorporating researcher action was included in a short discussion at the outset. This included highlighting the post Protocol Analysis debrief, as it was most likely that such a situation would occur during this interview approach rather than during a Cognitive Interview which was a post event reflection and recall. This is where a discussion of poor or unsafe student practice would occur, with the possibility of bringing it to the forefront of the interviewees mind or into their awareness. Whilst such an issue would have posed a significant dilemma, respect of my professional code directs that confidentiality be

overridden in the face of poor or substandard practice (Nursing and Midwifery Council and NMC 2008) in an effort to maintain patient safety. During the interviews one episode of poor or unsafe practice was identified by an interviewee, and a fail grade was awarded to that assessment artefact.

### 2.6.7 ETHICS OF PRACTITIONER RESEARCH

Practitioner researchers, as insiders, are able to explore issues which could not be investigated as thoroughly by an outsider researcher (Bridges 2001; Bridges 2009; Chenail 2011). Being part of the research milieu enables the researcher to understand the context within which interviewees exist, appreciating their lived experiences. I am a member of the academic community undertaking marking and grading of student written essays and as such have an understanding of some of the complexities involved. Fox et al., (2007) assert that practitioner research is a useful vantage point from which to examine and question the very nature of their own practice and that of colleagues. If a researcher occupies the same time, space, place and shared interests with the interviewees, a deeper understanding of the discourses, experiences and social realities of this shared community may evolve (Coupal 2005).

A disadvantage of practitioner research is over familiarity with the area, which could result in overlooking certain issues or even providing justifications for given situations. Bridges (2001) does contend that outsider researchers may be able to bring a better understanding of the research context, from a position aligned with no vested interests in the outcome. Outsider researchers, with limited knowledge of the context in which practitioners exist may misinterpret local nuances in practice situations (Thomson and Gunter 2011).

Insider researchers have the potential to impact upon the autonomy of interviewees who may feel obliged to participate. Previous or on-going positive or negative relationship histories may affect the freedom felt by the potential research participant to give consent (Nolen and Putten 2007). Respect for the ethical principle of autonomy could be considered to be

weakened, although a respect for the exercise of free will by potential interviewees was maintained (see Appendix 5)

Practitioner researchers have the potential to respect the ethical principle of beneficence (Gillon 1994) in relation to their own practice environment. In the current study I wish to explore the complexities of marking and grading to identify practical issues which could simplify the practice, thereby 'doing good'. This is seen as a benefit for interviewees to participate in the research process. Practitioner research enables the closer examination of the shared understandings of the local social practice of marking and grading, acknowledging the community of practice perspective.

## **APPROACH FOR DATA ANALYSIS**

### **2.7 INITIAL THOUGHTS ON FINDINGS**

The findings are concentrated on illuminating experiences of lecturers relating to marking and grading practice and is explored in Chapters 4 and 6. The creation of this research approach was developed to truly communicate Heidegger's view of phenomenology as the practice of "letting things show themselves" (Smith 2011). Researchers of this tradition acknowledge that they will need to unravel the data retrieved, to deduce meaning from their observations. In reality, this resulted in meaning being assembled from data gathered via recording the phenomena. As Crotty (1998) suggests, meaning does not exist independently in an object, but is realised by the individual perceiving the object at a particular moment. In relation to PA and CI methods, meaning was constructed from interpreting verbal data, constructing themes and classifying the data as I saw fit. This process was iterative, involving the researcher reconstructing meaning as new perspectives emerged through an on-going immersion in the data.

### **2.8 APPROACH FOR ANALYSING INTERVIEW DATA**

At the outset of the data analysis process there are a number of decisions that needed to be made to facilitate working with the available data in the most efficient way possible. As a novice researcher it was important to



carefully organise the data to facilitate an accurate audit of decisions made in the analytical process. After consideration of available options which were to use pen and paper or computer software for organising and analysing my data, a decision was made to utilise an available programme, NVIVO (Bazeley and Jackson 2013). All audio of the interviews and transcripts were imported into the programme to enable the process of analysis to begin.

A modified version of framework analysis was employed to analyse data from the CI interviews. This seemed to be a reasonable approach given I had attended training in the interviewing method at the centre which was known for the development of framework analysis (NatCen 2013). I had had the opportunity whilst there to sample framework analysis in its simplest form, is using the questions to develop a framework. The existence of an interview protocol, including a list of questions with accompanying probes for use with each participant, could be incorporated into a structure mimicking training received in framework development. To use this approach all questions and probes needed to be identified in each interviewees transcript with the answers from these questions and probes being collated for each interviewee. This approach would allow analysis by each question and accompanying probe for all interviewees. It should then be possible to group questions and responses into more manageable units or cluster into themes. Final decisions on the major and then sub themes all emanate from the original questions. This meant that the analysis would take a more deductive approach at the outset, using the questions as the sub context for exploring responses from the interviewees.

However this was not possible for two reasons: 1) interviewer style; 2) underpinning researcher ontology and epistemology alongside the research methodology. In relation to the first reason, prior to commencing this research I had had many years of experience at interviewing - for a variety of reasons and in a number of contexts, none of these as a researcher. The skills are transferrable and I found abilities I developed as a practitioner in clinical practice were soon put to good use as a manager undertaking recruitment interviews (following some training), and then again transferred to interviewing potential students for an academic programme. A focus on a

hermeneutic phenomenological enquiry, where I set out to discover the lived experiences of lecturers marking and grading student written assessment artefacts, fits with the second reason. I did not need to stick rigidly to the interview protocol. Instead I found I was able to actively listen to the interviewee in front of me and adapt my protocol as necessary to continue to encourage the conversation to flow in the most natural way. This reflects the chosen methodological approach of exploring interviewees lived experiences. So whilst I ensured that all of the items which I had felt at the beginning I needed to gather information on were covered, this was not through steadfast adherence to the protocol I developed at the outset. Completion of one interview to pilot the protocol, did not highlight this issue. Although it did result in two minor amendments to the questions, to ensure that I was able to gather the information on a range of issues I was expecting the questions to elicit.

### 2.8.1 EMPLOYING FRAMEWORK ANALYSIS

Analysis of the data did not commence until after completion of the data collection phase, which for some of the data meant an entire academic year had elapsed since the interviews. When the analysis phase commenced, on listening to the interviews it quickly became clear that inputting data into the framework developed around the questions was not feasible. However, this produced the most appropriate outcome, enabling close alignment with the research methodology and analysing the data set inductively. The analysis, in brief, therefore took the following form:

1. An initial listen to the recordings to check the transcription (or transcribe as necessary).
2. Review the list of questions to facilitate review of the transcripts.
3. Review and read each transcript, beginning to identify a number of themes. In NVIVO (which was used to manage the data and facilitate the process of organisation and analysis) these are called Nodes.
4. To ensure rigour, Supervisor to review two transcripts and to undertake an independent process of identifying themes. A

comparison of the same two transcripts and my identified themes took place, verifying the themes to which the data had been allocated.

5. Data in each of the themes (Nodes) that have been developed was reviewed, to establish its best fit (i.e. each Node in NVIVO). What became apparent is that there was some similarity with a number of themes (Nodes) and these were then collapsed into larger segments of data themes (Classifications in NVIVO).
6. Essentially this means now there is a list of sub themes and then major themes for the data set. All data for each major theme was examined against other major themes to ensure that there is no duplication and for best fit (though this should already have been achieved in step 4, above).
7. Data for each major theme with its integral sub themes is reported in the findings Chapters (4 and 6).

## 2.8.2 CI CODING AND THEME DEVELOPMENT

Coding of the data initially occurred by reading each CI transcript whilst simultaneously listening to the related audio recording. Performing these tasks concurrently facilitated a rapid immersion in the data along with providing an opportunity to correct any misheard words or phrases, making amendments to audio transcripts as necessary. This stage of the process took a large proportion of time, painstakingly reading and listening, however it had the desired effect of returning me to the context of the interview. On listening to the recordings in the initial stages it was difficult to decide what was of enough importance to note. As my confidence grew and understanding of the data developed, more interesting aspects of the data began to be noted.

The first phase of analysis was to code the data by identifying and grouping information together into themes as I saw them. Coding is accomplished by identifying passages, sentences or sections of text that appear to be important and then to highlight these and save them in a 'Node'. NVIVO uses its own terminology for handling and organising imported data, the term Node is used in the software to identify data which has been allocated to a particular code. Nodes are essentially the codes which became the themes

within my data, after reviewing the data set, a large number of themes were generated. Using NVIVO for the purposes of coding facilitated me in being able to review the list of themes as it grew. From the initial 10 themes identified in the first 2 interviews, the set of themes continued to expand resulting in 58 themes being identified by the end of the process. This list of themes was developed from across the range of sources derived from the 15 interviewees.

The next step was to review the individual themes, to identify and remove any duplicates. The occurrence of duplicates was a consequence of listening to the audio and reading transcripts on different days and resulted in devising some themes which on closer examination were not dissimilar from those previously created. At this point the number of themes was reduced to 55 (Figure 3). The next step was to then review the data that had been allocated to the duplicate themes to ensure that the new theme allocation was appropriate, by providing the best description for the data. NVIVO worked very well for this as I was able to collect data for each theme and then see it on the screen, either within the programme itself or to put it into a word document. Each of the similar themes were examined in this way, resulting in an initial reduction to 53 themes, whilst excluding two themes, treating these as outliers as they related to the interview method itself.

This list of themes was rather overwhelming, even when placed in separate text boxes again in NVIVO (Figure 3). To aid organisation Bazeley and Jackson (2013) advocate the utilisation of the model feature in NVIVO which would facilitate the next phase of data analysis, comprising of applying some sense of order to the themes.

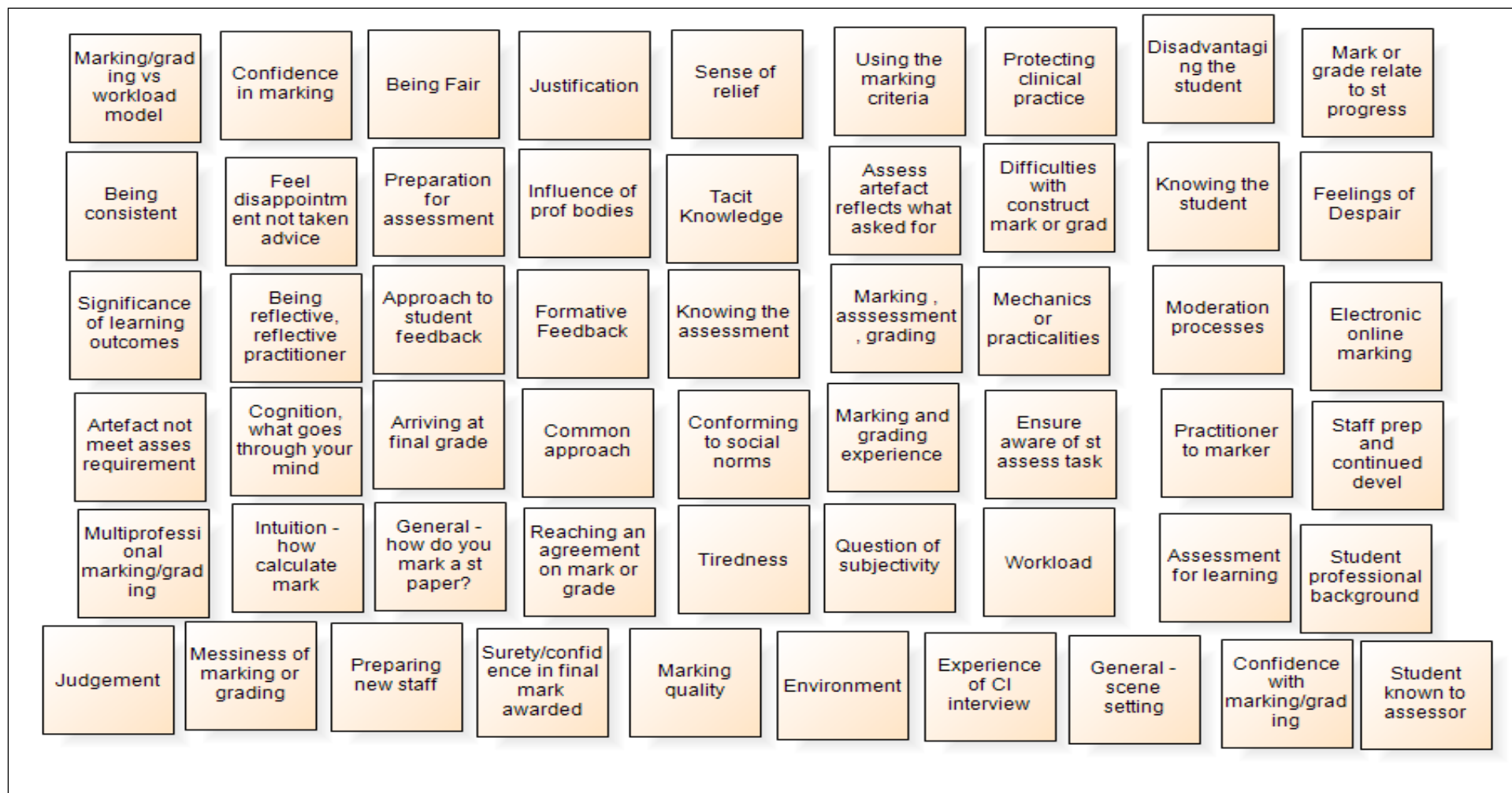


Figure 2 Original Static Model

The following phase involved the process of identifying how to work with this large number of themes. Rather than being confronted with a static list, NVIVO enabled each one of these themes to be located in text box which could be moved to any point on a page, which is a dynamic model in NVIVO. Each of the 55 text boxes, including two eventually identified as outliers, were then moved around the model until all the themes were grouped with what appeared to be like items or categories. Once this was completed, each category was reviewed and a title given to each which in some way summarised the subject of the themes. The model then enabled the data to be viewed as a set of major themes with a number of sub themes contained within them (Figure 4).

At this stage there remained some sub themes which appeared to be very similar. These were subjected to a more detailed review which included appraising all data which was attributable to these sub themes. NVIVO has the capacity to gather the data for each major theme and sub theme contained within it, whilst maintaining a link that is attributable to the individual interviewee from which it originated. In this way a fairly simple review process could take place as a first step in going back to the original data to check the soundness of the themes as shaped in these initial attempts. This process forms the beginning of step 3 in a modified Framework Analysis method of working with my data, Ward et al (2013) details this as a re-application of themes to virgin data. Within this current context the data was searched to retrieve extracts which had been coded at a particular sub theme, these were read carefully to identify meaning beyond the broad theme to which it had been allocated. If meaning could not be identified, the virgin data was reviewed at this point to try and establish more of the context beyond the individual excerpt. If it became clear that the sub theme allocation was inappropriate, the extract was re-coded, by being allocated to another major theme or sub theme at this point. This ensured that none of the original selected extracts were left outside a sub theme.

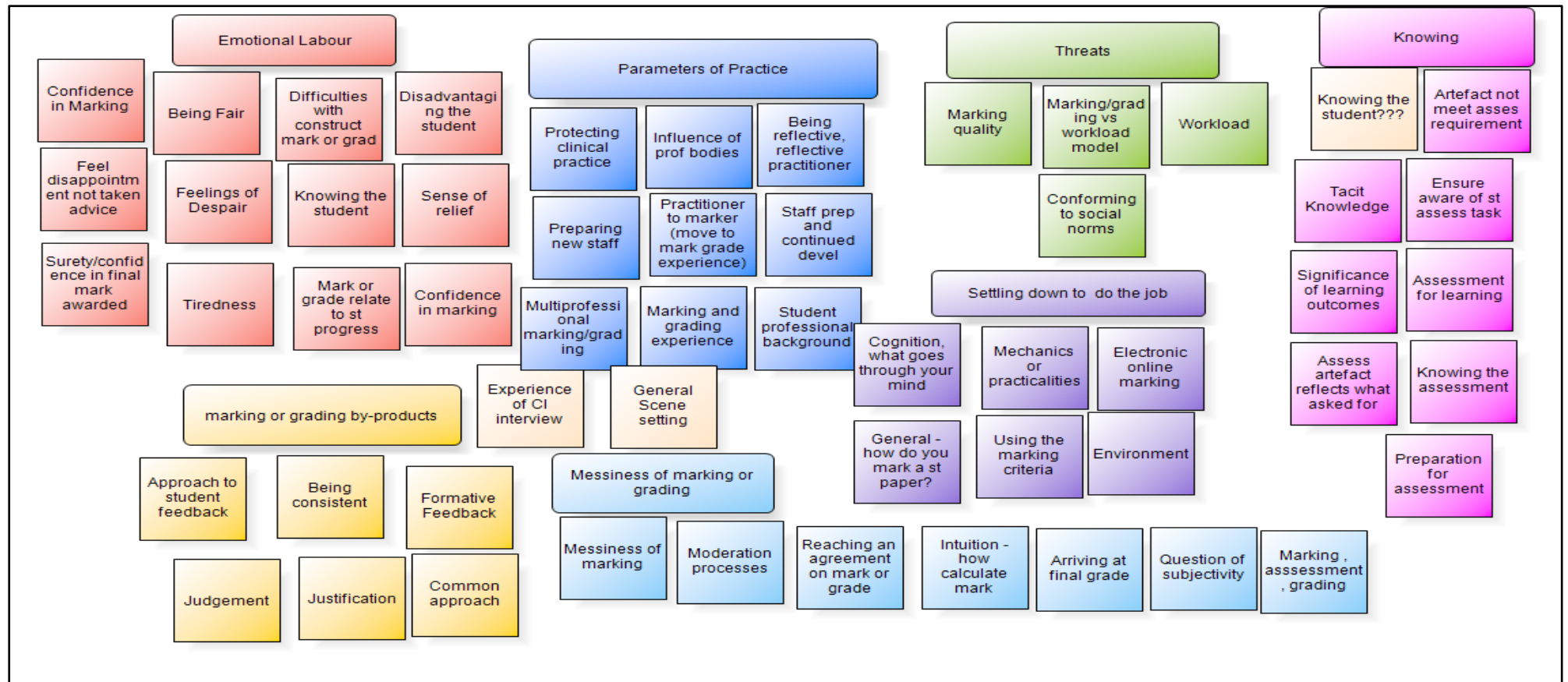


Figure 3 Original Major Themes and Sub Themes

Completion of this process resulted in some remodelling of the major themes and sub themes. This resulted in establishing 5 major themes with embedded sub themes. A record of these changes was kept in NVIVO by the creation of 'static models' showing the final make up of major themes and sub themes (Figure 5) which facilitated an audit trail through the analysis process.

## 2.9 IDENTIFYING AN ANALYTICAL APPROACH FOR PA INTERVIEWS

The aim for analysis of data gathered from PA interviews was to listen to each of the 11 recordings and to simply identify themes. Closer review of the literature related to data analysis (Van Manen 1990; Fleming et al. 2003; Jacelon and O'Dell 2005; Halcomb and Davidson 2006; Bazeley 2009; Smith et al. 2009; Bazeley 2013) establishes that this initial approach could be viewed as requiring increased rigour relating to dependability, credibility, transferability and confirmability. Employing a recognised tool for conducting data analysis enables the presentation of findings which are auditable by another researcher. At the beginning of this report, I firmly establish my pre-existing knowledge and connections to the phenomena which goes some way toward facilitating an understanding of my roles as researcher and research instrument (Mason 2002b). Through utilising researcher reflexivity it is possible to explore what I see as my intersection with the research from my impact on the research environment to my interpretation of the data. I acknowledge that as the researcher I constructed meaning from what was already present, with my interpretation of findings being seen as socially constructed, bounded both contextually and situationally. Thus the historical, cultural and contextual significance of the data collection cannot be ignored, with my selection of interview methods allowing meaning, which is no longer concealed, to be brought into view. Such clear explanation of the research context is provided to facilitate the reader identifying whether findings can be reliably transferred to another context which includes marking and grading. Finally, the use of an identified framework supports the verification of confirmability of how interpretations were reached in my study. Heideggarian



hermeneutic phenomenology (see 2.5.2 ) is seen as an appropriate approach to expose interviewees world of marking and grading, adding researcher reflexivity (see 2.5.4 and 8.3) clarifies how interpretation of the data, and construction and reconstruction of findings has been achieved.

Modified framework analysis was identified as the most appropriate method within the context of my research study, enabling a guided systematic approach to the exploration, analysis and presentation of findings for this research method. Alongside providing a common approach to data management across this research thesis, framework analysis also facilitated a transparent approach to understanding the outcomes of the analysis process.

## 2.10 PA CODING AND THEME DEVELOPMENT

All PA interviews were recorded; I listened to each of the 11 interviews in turn, to search for themes. Verbatim transcripts were created although Bazeley (2013) and Creswell (2012) assert that this is not necessary. They suggest that the researcher can listen to the recordings to become immersed in the data, facilitating the search for and identification of themes.

One issue to consider at the outset of the analysis process was that this research was not focused on traditional problem solving. In the research literature PA interviews were conventionally used to gather information on approaches to participant selection of responses in the review of website usability and design (Benbunan-Fich 2001). The aim of this research is not the search for this kind of specific data. As this is the case there are no apriori conditions to satisfy.

A step by step process was completed applying the principles of a modified version of framework analysis, in the same methodical way as this had been applied to the CI data. When all the themes had been created, a global view of them was taken which enabled an overview resulting in the identification of duplicate or like themes. The PA interviews yielded fewer sub themes because the majority of the interview was focused on the engagement and

interaction of the interviewee with the assessment artefact. The final list of themes resulted in 27 sub themes which was reduced from an initial total of 48. Figure 11 PA Initial Coding shows a tree model created by NVIVO. This data analysis software was utilised for organising and displaying my data in the constructed major and sub themes.

To reduce prejudice, as per a Gadamerian perspective (Fleming et al. 2003), a random selection of interviews were listened to by two Supervisors. This element of checking was in a bid to ensure not only the credibility and dependability of the final research product, but to increase trustworthiness. Utilising Supervisors to listen to interview recordings, and confirm or reject the themes I identified as one aspect of the triangulation. Member checking was used as another method to verify the application of themes in the production of this thesis (see 5.2.6).

## 2.11 DATA COLLECTION LIMITATIONS

The biggest technical limitation identified with this study was related to issues with the clarity of recordings for transcription purposes. Three PA interviews were unusable, unfortunately the positioning of the directional microphone on the desk, next to the keyboard was inappropriate. Quietly spoken interviewee voices were obliterated by keystrokes from simultaneous composing of feedback for students. This was combined with scrolling up and down the assessment artefact (which in every case was presented electronically) to try and identify elements interviewees felt they had to read at that moment. In hindsight a microphone which could be pinned to the lapel of the interviewee would probably have provided the best sound quality.

In the case of the CI sections of several interviews are difficult to hear due to the external environment. Despite asking if somewhere quiet could be found to conduct the interview, restrictions were encountered. Much office space is shared and difficulties with negotiating room booking systems meant that on two or three occasion's interviews took place in a more public environment. Regardless of being in what appeared to be a quiet corner, general

environmental background noise again impacted upon the quality of the final recording.

Unfortunately, as a result of the difficulties highlighted above, a small proportion of interviewee stories had to be abandoned and are therefore not included in the final analysis. Although not ideal, Chapters 4 and 6 demonstrate that findings reported share commonalities in the topic areas explored and conclusions drawn by those interviewee stories that could be clearly heard and therefore transcribed. As a 'novice' researcher my sound check consisted of looking at the front of the digital recorder when individuals or I were speaking to see if there was a rise and then fall in the sound level indicator. Instead an appropriate sound check should have been carried out with each interviewee, which would have included an opportunity to playback sounds of interviewees speaking.

## 2.12 SUMMARY

A study of the philosophy of research has highlighted that rather than simply focusing on a method for retrieving data, it is imperative to become cognisant of my own ideas relating to ontology and epistemology. A conscious appreciation of ontology and epistemology has led to increased confidence that the data gathered is a reflection of the true essence of nature. This is not withstanding forgoing discussions on the importance of viewing findings as produced in a particular context and consequently as being historically and culturally mediated. Thus the methods selected to investigate the nature of marking and grading become congruent with this ontological and epistemological approach. It is these two strands, my orientations as a researcher rather than the methods which will produce specific forms of knowledge (Morgan and Smircich 1980) providing an answer to the research questions. Outcomes of this research, by application of the philosophy explored here are reflected in the methodology and also the methods selected (Chapters 3 and 5) and will direct future experience of marking and grading practices of Novice and Experienced lecturers in Higher Education.

Research to inform this thesis was completed by utilising two separate but potentially complementary methods. Both methods adhered closely to the tenants of phenomenology, enabling me to gain close contact with the world or lived experience of lecturers and the task of marking and grading. The nuances of each method Chapters (3 and 5) and rationale for their selection prior to the related findings are explored in the forthcoming Chapters (4 and 6).

# Chapter 3 COGNITIVE INTERVIEWING METHOD

*No man can reveal to you nothing but that which already lies half-asleep in the dawning of your knowledge.*

Khalil Gibran

This chapter is concerned with exploring the first method of data collection utilised in my research study, Cognitive Interviewing (CI). The background of this method is explored with identification of some of the benefits of this method, whilst acknowledging some of the limitations documented in the literature. A rationale for its choice leads to an exploration of interviewer preparation, interview process, and interviewee selection. The approach taken to analyse data gathered in this study has been explored in Chapter 2, however, toward the conclusion of this current chapter the approach taken to provide assurance of the credibility and dependability of findings is presented. This chapter concludes with a summary of CI method of data collection. **WHAT IS COGNITIVE INTERVIEWING?**

Cognitive Interviewing was originally developed in the United States by two psychologists, Fisher and Geiselman in response to their investigation of police interviewing techniques. They identified that the standard of interviewing of eyewitnesses was variable and times inadequate. They considered that the importance of obtaining comprehensive and truthful information from eyewitnesses, in cases of being witness to a crime, could not be overstated. In instances where accurate evidence was not generated this was more likely to lead to a miscarriage of justice, resulting in an innocent individual facing prosecution (Fisher 1995). An initial literature review by Geiselman et al. (1986) established that it was not common place for police investigators to be given any training on how to conduct an effective interview. It appeared, in the main that investigators relied upon using interview techniques learned in their initial training, as well as techniques learned in through on the job training or observation as a police officer, and also via intuition. Given the importance of the retrieval of accurate statements, it appears appropriate to explore methods used to access the memories of eyewitnesses. Cognitive Interviewing (CI) employs

current theories of memory retrieval from the field of cognitive psychology (Atkinson and Shiffrin 1968; Baddeley 1997) to increase the completeness of witness reports i.e. improve the outcomes of police interview procedures (Fisher et al. 1989). Geiselman et al. (1986) conducted research to evaluate the efficacy of this innovative technique. Indeed Fisher et al. (1989) demonstrated that the Cognitive Interview was able to reliably enhance memory, with witnesses recalling events with improved accuracy, and advocated its use for a variety of investigative interviews.

Development of the CI technique and concomitant set of instructions are based on generally accepted scientific principles of memory (Fisher et al. 1989; Tourangeau 1999). The technique is used to prompt effective retrieval of memories of a specified event and it has been demonstrated to elicit 25 – 35% more information than a standard police interview (Geiselman et al. 1986). More importantly this is achieved without producing an increase in the amount of inaccurate or incorrect information being conveyed. The Cognitive Interview approach is utilized in survey research to pre-test questions on survey type questionnaires (Memon and Bull 1991; Memon and Higham 1999; Drennan 2003). This is primarily to establish the usability of questionnaires, exploring the ease and accuracy with which interviewees could complete a questionnaire. Survey questionnaires can be paper based or computer generated, with Cognitive Interviewing Interviews being used to test the questions on the questionnaire. The purpose being to establish whether or not intended interviewees can understand what is required of them when answering the questionnaire. NATCen (National Centre for Social Research) use this research approach to validate their questionnaires prior to using them in pilot research and then in main research studies (NatCen 2013). For instance, the National Census is conducted by questionnaire and the questions are pre-tested to ensure they are understandable along with being able to elicit the information required (Drennan 2003; NatCen 2013). Questions should be unambiguous with no or limited possibility of directing interviewees away from their intended focus. Using Cognitive Interviewing Interviews provides the opportunity for researchers to test each question, whether answered or skipped and to

explore with the research interviewee the reasons these actions occurred (Drennan 2003). This also means any lexical problems can be clarified, at the time, by the researcher and the interviewee.

This technique facilitates the detection of issues that may remain hidden with other types of data collection (Tourangeau 1999) for example through the use of questionnaires or pure observation. Retrieval cues and mnemonics are used to uncover what might not be discovered by other data recovery methods. These mnemonics were cited by Geiselman et al. (1986) and consist of four processes, two of which are designed to increase the overlap between the reporting situation and the event – reinstate the environment if possible, and to encourage the interviewee to report everything. The final two processes are related to how the event is reported, interviewees are encouraged to explore the event in different orders and from different perspectives in the hope of revealing information with a depth and breadth not achieved through ordinary approaches to storytelling.

### 3.2 PROCEDURE

A particular procedure is instigated to utilise the 4 principles of memory retrieval advocated by Fisher et al. (1989). These principles are ‘Event – Interview Similarity’; ‘Focused Retrieval’; ‘Extensive Retrieval’; and ‘Interviewees Compatible Questioning’, and they will now be explored in turn.

It is suggested that memory can be enhanced by reinstating the psychological environment at the interview which was similar to the event (Geiselman et al. 1986) i.e. ‘Event-Interview Similarity’. The event in the case of my research was marking and grading completed by the interviewee. It is proposed that using retrieval cues with the interviewees, which have an overlap in relation to the features of the ‘to be remembered event’, would arouse their memory. This was achieved by asking the interviewee to describe aspects of the event and environment e.g. what the marking and grading was, when the marking and grading took place and where this was undertaken i.e. home, work or elsewhere?

'Focused Retrieval' is the second principle of memory recovery and is based on the premise that information about the event may well be encoded in different ways. There is evidence to suggest that memory retrieval takes a great deal of 'focused effort'. To encourage the interviewee to maintain this effort the interviewer must actively listen and use probes as and when necessary to increase or continue information retrieval. Use of different retrieval cues will be required to enable access to the different encoded paths, aiding accurate recovery of the 'to be remembered event' ( Craik and Lockhart 1972; Geiselman et al. 1986; Tourangeau 1999). In relation to my research this took the form of using probes (See Cognitive Interview Protocol Appendix 6).

The third principle is to execute 'Extensive Retrieval'. Evidence suggests that encouraging the interviewee to make a number of attempts at retrieving the event, rather than a single attempt will facilitate more in depth recovery of information (Geiselman et al. 1986). This takes the perspective of leading interviewees to closely examine general aspects or mechanics of marking and grading prior to focusing on a single assessment artefact. Interviewees are then instructed to move from this close focus to expanding back out to thinking about the impact of evaluation of this particular artefact and their view of marking and grading as a whole. Interviewees are led from the macro i.e. marking in general to the micro, an individual script, back out to the macro thinking of the module as a whole in relation to marking and grading of student artefacts.

The fourth principle is 'Interviewee Compatible Questioning'. It is thought that every interviewee stores and organises memories in different ways. In order to aid the interviewee to retrieve information about the event, the interviewer needs to be responsive to the physical and verbal cues which would illustrate the interviewees understanding and interpretation of the questions. It is important therefore to adapt the questions to enable the interviewee to explore memories in a way which matches their needs. Each interviewee will have different needs and as such the interviewer should adjust the questions to match the situation accordingly. In my study this is evident from reviewing the interview transcripts where questions were phrased differently for



different interviewees and the choice of probes was a decision made at the time of the interview. Probe selection was based on what was a good fit for the interview situation. An additional aspect of the CI is the use of a number of mnemonics (Fisher et al. 1989) which facilitate the remembering of specific information or may assist with eliciting fragments of information which can then be explored through further questioning.

All communication is directed toward improving memory retrieval and enabling the interviewee to communicate these to the interviewer. The aim is to support the interviewee to convert recollections into a response; to ensure that the interviewee remains focused on retrieving and communicating information relevant to the needs of the interviewer; direct communication to facilitate the interviewer's comprehension of interviewee responses; and those used to engender interviewer understanding of the interviewees psychological needs. In my study the mnemonics used during the interview process were based around probes designed into the CI Protocol (Appendix 6). However, in the spirit of 'Interviewee Compatible Questioning', these were used only when deemed necessary to enable clarification of any mystifying aspects.

Effective CI is focused on guiding the interviewee toward an enhanced level of memory retrieval. This was achieved in my study by asking the interviewee to think about the concept of marking and grading, in general, of written student submissions. They were then implored to talk about their processes and finally to focus on a specific module, for which they had completed marking and grading sometime within the last calendar year. They were then to think about the marking for that module as a whole and then to concentrate on a single artefact, one which stood out in their memory.

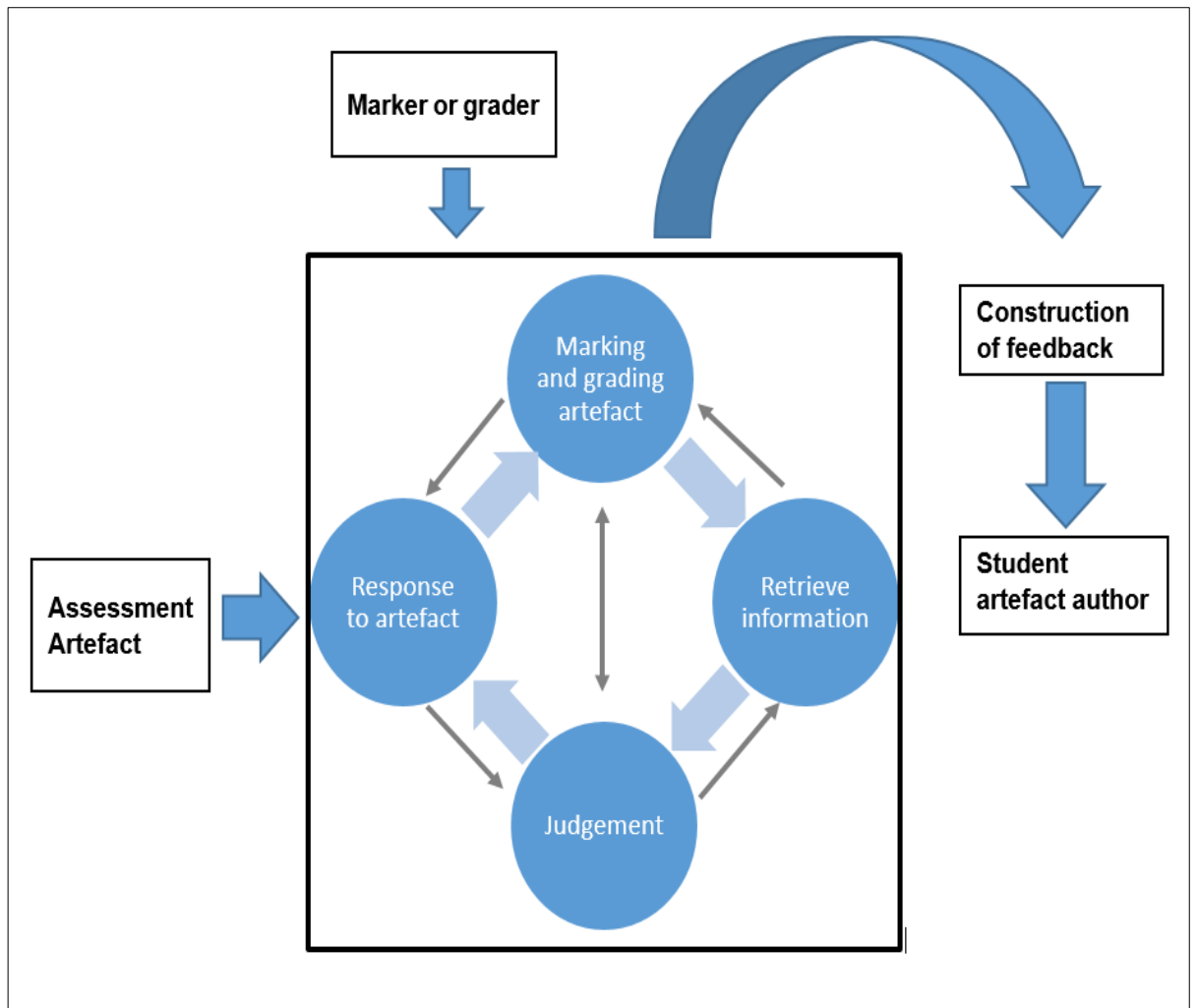
### 3.2.1 ADVANTAGES

Cognitive Interviewing Interviews were selected as a method for data collection as a face to face intervention rather than selecting a method which required no interpersonal connection, for instance when using a questionnaire. In order to gain in depth qualitative data, facilitating the

lecturer providing context to the process of marking and grading, a CI was considered to be an effective approach. Face to face interviews enable researchers to identify and explore any interpretation or comprehension difficulties in relation to questions, to explore at the time of the interviewees retrieval processes, exploring their cognitive understanding or memories of the events.

Use of cognitive theory will aid in the understanding of information processing, problem solving, and reasoning in relation to the task of marking and grading. In my research the data gathered was relevant to gaining deeper understanding of how lecturers responded to artefacts requiring marking and grading. Using an interview enabled the interviewee to answer more explicitly, with the interviewer being able to clarify any ambiguities (Tourangeau et al. 2000). Also this technique ensured that there was no break in flow of the event i.e. marking and grading in the case of this research. CI focuses on the mental processes used in making judgements and in the case of this research it carefully explored these in relation to marking and grading. Interviewees were asked to notice how they responded to the marking and grading task, exploring carefully a series of complex cognitive processes.

Tourangeau et al. (2000) developed a model for comprehension, which has been adapted for use in this study. Interviewees are required to retrieve from memory relevant information related to the written assessment artefact being considered, including recalling all processes used to come to a judgement about the artefact illuminating estimation and response processes. These are all mapped to the potential response options for the artefact e.g. excellent, very good, good, or fail. Marking and grading is not linear in terms of engagement with the process and there will inevitably be some reciprocity involved in attaining a final judgement as to the value of the artefact. This is depicted in the model below:



**Figure 4 Marking and Grading Processes**

*In Figure 4 Marking and Grading Processes, the lecturer has an artefact to mark and grade, they read it, retrieve information about its quality or otherwise, thus responding to the work before them. At some point they will need to make a judgement. The model illustrates the movement between these elements, which is not cyclical but involves returning to elements of the process until a decision is made.*

A further advantage of this approach is the ability to explore the interviewees individual definitions of the marking and grading process. This was achieved by using direct questions (where required) which fitted with the retrieval mnemonics of Geiselman et al. (1986).

### 3.2.2 DISADVANTAGES

One of the main disadvantages of CI is that information is retrieved as a post event reflection. Memory and recall are known to be affected by time which has elapsed since the event ((Tourangeau et al. 2000). Memory decays over time, and this has the potential to have a negative impact upon interviewees ability to recall progression of the event, including all nuances accurately. It is therefore not unusual for interviewees to forget what they were thinking at the time of the event, including not remembering what had an influence on their actions or reaction. Frequently occurring events are also more likely to be difficult for interviewees to remember accurately in relation to when it occurred. This is because the frequency makes the occurrence more routine and commonplace with no defining features, clear recall becomes all the more challenging. In essence the topography of each marking and grading event is similar and is thus a common task to be completed by lecturers.

My curiosity in this commonly undertaken task extends to all the small and potentially insignificant or unremarkable nuances which occur during the process of coming to a judgement about an assessment artefact when providing a mark or grade. The danger here is that this type of information becomes more easily blurred over time, and more difficult to recover.

Marking and grading is a common event in the working life of a lecturer and therefore mundane unremarkable nuances are expected to easily blend into the generic memory, distinctly affecting the possibility of accurate retrieval.

Current theories examining the effectiveness of CI are based on measuring the gain in the number of correct statements elicited by interviewees or witnesses (Geiselman et al. 1986; Memon and Bull 1991; Saywitz et al. 1992) . Within the research reported here there can be no appropriate measure to establish the effectiveness of this technique as interviewees are not expected to provide correct or incorrect answers to questions. They are simply required to relay their experience of marking and grading in as much detail as possible, to be achieved by re-instatement of the context within which the practise took place. This fits with the hermeneutic phenomenological approach upon which my research is based.

Whilst it is acknowledged that some memory loss is inevitable, CI tries to diminish this by the use of an open questioning style, following an event along with the use of specially selected probes and mnemonics (Fisher et al. 1989). In my research, interviews did not take place immediately after the event. In some cases a great deal of time had elapsed from completion of the marking and grading that an interviewee was trying to remember. As it is accepted that memory retrieval decreases as time following an event increases, for these interviewees ability to recall their actions, reactions and nuances about marking and grading will be impaired. This had the potential to impact upon the quality of information retrieved from interviewees.

### 3.2.3 COGNITIVE INTERVIEW PROTOCOL

An issue worthy of consideration with CI is the increased potential for researcher bias as the interviewer can ask leading questions, with further leading follow-up probes, to increase the depth and breadth of information retrieved. As a consequence of this there is a capacity for retrieval of artificial findings, generated through the interviewer using inappropriate techniques which threatens the collection of authentic data (Tourangeau et al. 2000). To combat this, I developed a CI Protocol (APPENDIX 6) which included a guide to questions and potential probes to be used which was submitted with the application to my home institution University Research Ethics Committee, for which approval was granted (APPENDIX 1). As this study also involved gathering data at 3 external institutions, the CI Protocol was reviewed and given ethical clearance by each prior to the commencement of data collection. Whilst no formal pilot study was undertaken, a 'test' interview was carried out with my supervisor prior to use in the field.

Memon (1999) identifies a particular issue in relation to the interviewer when undertaking this type of interview. There is an increased demand on the cognitive load of the interviewer in comparison to other types of qualitative interview. This is because of the requirement to engage in attentive active listening, facilitating application of potential probes effectively to increase access to cognitive processes related to the experience being explored.

### 3.2.4 INTERVIEWER TRAINING

Memon (1999) advocates that any interviewer using this technique undergo training prior to utilising CI for gathering data. They recommend that this training should be 2 days in length. I undertook this training at the National Centre for Social Research (NatCen 2013) who have extensive experience in CI, although this is in relation to survey question testing. I completed the training prior to commencing any CI interviews, this increased my knowledge and confidence in the principles and application of the techniques. Whilst Fisher et al. (1989) and Geiselman et al. (1986) describe how experienced and inexperienced police detectives are able to use the technique with little prior training. CI was found to be effective at eliciting increased levels of information including attention to completeness of information recalled, achieved through the application of the four general memory retrieval techniques. Although no evidence exists which can confirm CI efficacy when there is a long delay between the experience and occurrence of the interview and its effect on data retrieval (Memon et al. 2010).

### 3.2.5 CONDUCTING COGNITIVE INTERVIEWING INTERVIEWS

The fundamental purpose of this research was to identify and explore, through in-depth semi-structured Cognitive Interviewing interviews, how lecturers approached their marking and grading, setting out what they did. This technique was used to facilitate the retrieval of data and this was achieved through, a more researcher led method than is used in traditional qualitative research interviews. Completion of the interview involved the interviewee participating in a post-performance reflection on an authentic task, the task of marking and grading. Interviewees were led through the process, initially talking generally about their feelings and or approaches to marking and grading. The interviewees were directed into reflecting on marking and grading by focusing explicitly on their most recent experience; this facilitated the gathering of detailed data on the subject. The process was guided by use of mnemonics as directed by Geiselman et al. (1986) and Fisher et al. (1989). The accompanying CI Protocol was designed following training and was based on the format used by NATCen (2013).

For conducting the interview a quiet environment free from interruptions was most important. This proved to be unachievable in all cases; many academics did not have access to their own space as they work in shared offices at their academic institution. Discussion with the interviewee took place before data collection began as it was important to arrange the most appropriate place for the interview. Ensuring that the interviewee felt comfortable and able to disclose freely any information about their approach to marking and grading whilst maintaining their anonymity was important. To this end flexibility was required to complete interviews in settings chosen by the interviewee, which were both on and off campus. In order to maintain safety, I ensured that my appointment schedule was known by another person including estimated time of return.

### 3.2.6 SELECTION OF INTERVIEWEES TO CONTRIBUTE

Literature reviews promoting the efficacy of CI give limited attention to the potential number and composition of interviewees for inclusion in a research sample (Beatty and Willis 2007). In the absence of such detailed information I developed inclusion and exclusion criteria which served as a guide to the composition of interviewees for this data collection method. The inclusion criteria included a requirement for the interviewee to have completed marking within the last calendar year. This was slightly at odds with my original proposal of selecting interviewees who had completed marking and grading within the current academic year. This standard was selected as it had a potential advantage of using CI to retrieve in-depth post event reflection by the interviewees. However, after final achievement of ethical approval at 4 institutions and timing of interviewees having access to marking and grading in the semester in which data collection took place, was not always feasible. Therefore identified techniques for CI were applied to aid interviewees' retrospective recollection of their approaches to marking and grading. Interviewees were asked to recall all procedures, processes, thoughts and feelings whilst marking and grading an identified set of student scripts, focusing specifically on the construction of a mark or grade for one or two written assessment artefacts.

### 3.3 SUMMARY

In summary, Cognitive Interviewing through questioning and follow up probes is a method that attempts to elicit and expose cognitive process i.e. the thoughts of the interviewee in relation to the question posed. When used to test survey questions it is not the answer to the survey question that is of interest to the interviewer, rather it is the thoughts and process that the question evoked i.e. was the question understood, did the answer given fit with what I originally intended or even were there any words or terms that the participant did not understand? Gaining an in-depth representation of these processes will provide greater assurance that when these questions are used in large scale surveys, they will produce limited amounts of misunderstanding, or misinterpretation. To date this has been the most frequent use for this tool (NatCen, 2013).

The Cognitive Interviewing technique has also been used extensively, in another area, to elicit in-depth information involving the interviewee remembering an event or events and associated thoughts and feelings. This area is in police interviewing and witness statement development (Geiselman et al. 1986; Fisher et al. 1989) to aid investigators in the detection of crime. This type of interview can appear to be more researcher led than traditional qualitative interviews; they are focused on improving recall in relation to an event, in my research the event was marking and grading written assessment artefacts. This focus enabled the gathering of specific data on the object of study. Consequently it is considered here to be a suitable technique to use in order to return interviewees to the situation in which they were undertaking marking and grading, to explore their approach in fine detail.



# Chapter 4 COGNITIVE INTERVIEW FINDINGS

*The obvious is that which is never seen until someone expresses it simply.*

Khalil Gibran

This chapter presents findings from the Cognitive Interview (CI) method of data collection and using modified framework analysis facilitated working with the data. The processes involved in analysis were explored in Chapter 2, facilitating a transparent approach to understanding the outcomes of the analysis process.

This chapter is then organised in the following way: inclusion of the research aim, questions and objectives, information about the research cohort.

Through close exploration of the data, lived experiences of interviewees' marking and grading will be illustrated before closer analysis is completed and connection to contemporary literature is made in Chapter 7.

## 4.1 RESEARCH AIM

To investigate the process of marking and grading from the perspective of the marker. To recognise good marking and grading practice, deriving a model of 'expert' practice, aiding the development and proficiency of novice markers.

This interview method proposes to uncover answers to two of the three thesis research questions (1 and 3), whilst attending to the research aim, and maintaining a connection with the research objectives.

### **Research Objectives:**

1. Examine assessment by exploring marking and grading practices.
2. Identify what are 'good' marking practices to develop a concept of 'expertise' in marking and grading.
3. Explore novice lecturer's thoughts on marking and grading.
4. Explore cognitive processes and extraneous influences on marking and grading practice.

### Research Questions:

1. What do lecturers believe they do when they are undertaking marking and grading?
2. What thoughts or cognitive processes occur in lecturers during the act of marking and grading?
3. What exists to ensure equity or consistency in professional or inter-professional marking and grading, this includes inter-rater reliability.

## 4.2 INTERVIEWEE KEY 1:

Table 3 CI Interviewees

<b>PSEUDONYM</b>	<b>EXPERIENCE in years</b>	<b>INSTITUTION (Current)</b>
Chris	20+	Pre 1992
Gerri	20+	Pre 1992
Kim	5-10	Post 1992
Eddi	20+	Pre 1992
Terri	5-10	Post 1992
Billie	5-10	Pre 1992
Lesli	5-10	Pre 1992
Mel	Less than 2	Post 1992
Danni	10-15	Post 1992
Harri	10-15	Pre 1992
Alex	15-20	Pre 1992
Francis	10-15	Pre 1992
Hilary	5 - 10	Pre 1992
Izzie	2 -5	Pre1992
Jamie	Less than 2	Post 1992

The table above provides some context for reading the interviewees' stories presented below.

## 4.3 EXPLORATION OF FINDINGS

To explore the data and draw out the findings, modified framework analysis was employed. The process taken is explained in detail in Chapter 2, but it enabled examination of the transcripts of each individual interviewee to elicit the production of codes which became sub themes. Finally sub themes were grouped together to form seven major themes, as I saw them, and are illustrated in Figure 5. The following findings are reporting on five of those major themes Figure 5 which are directed at answering the research questions: 4.3.1 Settling down to do the job, 4.3.2 Parameters of practice, 4.3.3 Messiness of marking and grading, 4.3.4 Knowing and 4.3.5 Marking and grading by-products. Theme number 6 is Emotional labour which has not been included for further analysis here, but will form the basis of future work. The final theme, Outliers will be explored in Chapter 8 in relation to researcher reflexivity.

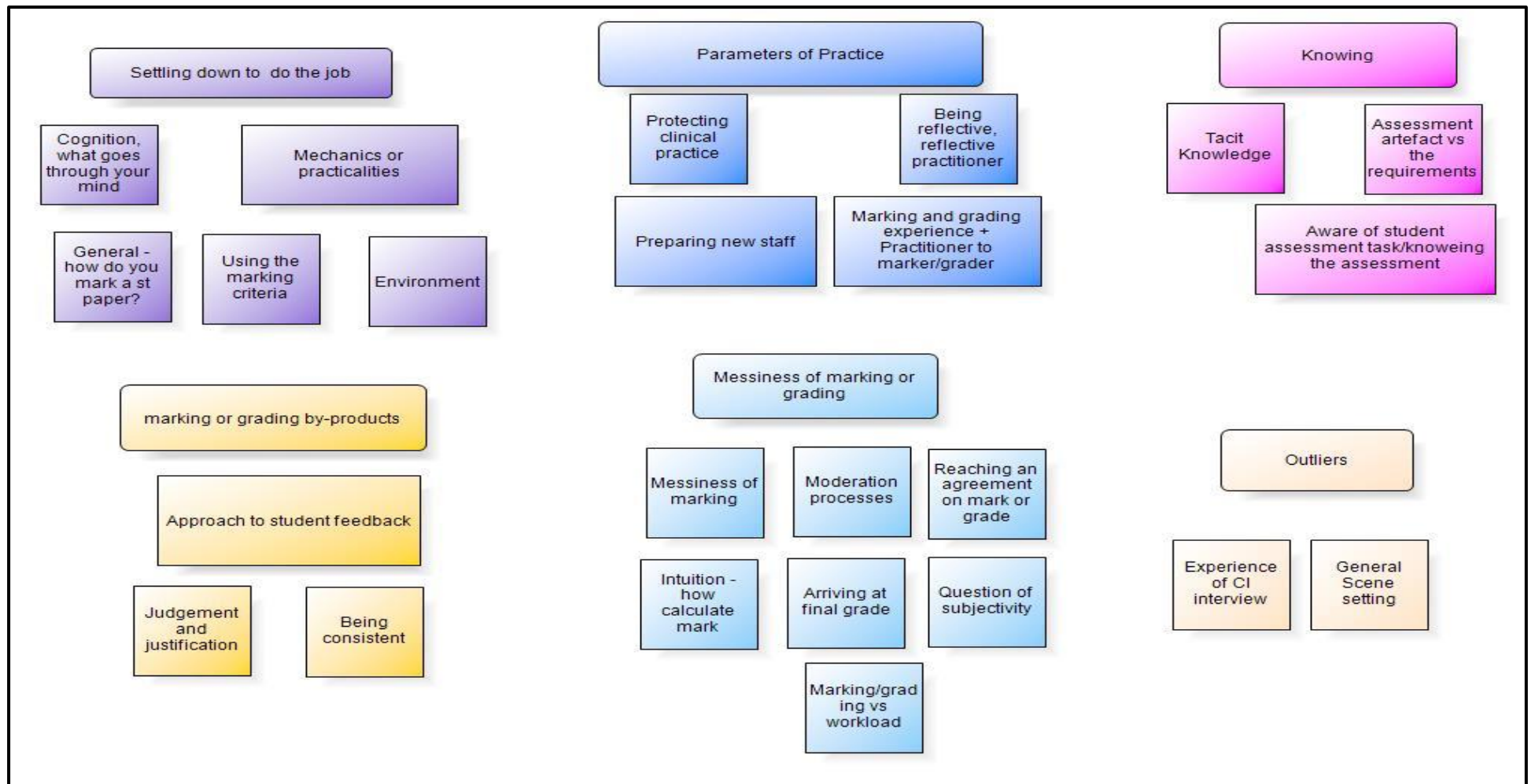


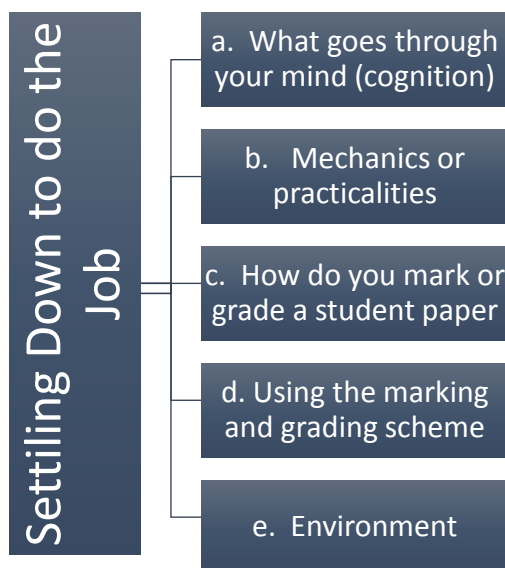
Figure 5 CI Final Major Themes and Sub Themes

No hierarchical order for presentation of the themes has been established, but each will include excerpts of data. This interview method yielded a large amount of rich data, presenting an in-depth illustration of the experiences of lecturers executing their marking and grading roles. It was not possible to include all examples of data for each sub theme. Data presented does not include examples from all 15 CI interviewees, instead the excerpts selected are those which appear to capture the essence of the theme to which it was allocated. The major themes are as follows: 4.3.1 Settling down to do the job, 4.3.2 Parameters of practice, 4.3.3 Messiness (Hunter and Smith 2007) of marking and grading, 4.3.4 Knowing, 4.3.5 Marking and grading by-products.

The final list of 22 sub themes was created excluding the two outliers, general scene setting and general experience of the Cognitive Interview. An identification of the major theme with its allocated sub themes is presented as a figure at the commencement of each of the following sections. This precedes exploration, explanation and critique of the findings from the CI interviews.

### 4.3.1 SETTling DOWN TO DO THE JOB

This first major theme is examined by interviewees and explores their approach to commencing the task of marking and grading that is before them. This major theme is examined through five sub themes. To provide context quotes from interviewees will be included along with a short introduction, then interpretations and concluding summary.



**Figure 6 CI Major Themes**

#### 4.3.1.a WHAT GOES THROUGH YOUR MIND (COGNITION)

All interviewees were asked about what they remember going through their minds when they are undertaking marking and grading of an assessment artefact. Interviewees were encouraged to voice what they remembered their initial thoughts to be. A range of responses to this question are reported here, which include lecturers' first impression upon being exposed to the assessment artefact, as they see it.

'Eddi' stated:

*"The first thing I thought was it doesn't look very good, it doesn't bode very well when there is a grammatical and a spelling error in the first sentence. That was the first thing that went through my mind"*

This illustrates that this interviewee quite quickly came to a judgement as to the quality of the work they were evaluating.

'Eddi' also voices another aspect of the initial thoughts, whilst this is related to workload it is worth exploring the impact of these thoughts:

*"OK it's, so it is about getting them done. The first thing is how can I get these finished in the time that is available cos sometimes there's flipping hundreds of them".*

These initial thoughts may have little influence on the evaluation processes to follow. At the outset of a task, having an impending sense of overload may mean that the task is completed with less attention than would be warranted, in order to meet agreed deadlines.

'Gerri' approaches the first contact with the written assessment artefacts wearing a very different set of lenses, specifically focusing on positive aspects:

*"I think what goes through my mind always is not a sense of excitement exactly, anticipation, what am I going to get here you know".*

It seems from these interviewees that the initial contact with the assessment artefact is important, in providing a frame of reference from within which to make their evaluation. 'Harri' is able to articulate this process:

*"I'd like to think that I keep an open mind I can't be absolutely sure that I do. But .... I know it's about first impressions but I have had papers that have been so good to start with and they've obviously just lost interest or actually they've not been as good as I thought they were, and so then you start to think differently ... I think, the trouble is you don't often get a bad one that gets better".*

This interviewee is exploring what having an open mind means to them and how this aids them in the marking and grading process.

These interviewees all express differing aspects of the thoughts that occur early on in coming into contact with the assessment artefact, much of which focuses on coming to an initial judgement or evaluation.

#### 4.3.1.b MECHANICS OR PRACTICALITIES

Interviewees were encouraged to discuss in a step by step way what they think they actually did, and in what order, to complete the task of marking and grading a written assessment artefact.

'Gerri' identifies a comprehensive process which they instigate before they can commence marking and grading of the assessment artefacts:

*"... we do all our marking on a particular schedule or sheet, a marking sheet which is roughly the same across all the programmes ... it is done anonymously. Which assignment, which module, which programme it is all the usual things ... I don't like to be marking and then having to type all these details I like to get all that set up so, I've got twenty scripts to mark I'll have twenty of these sitting on my computer ready to fill in. Then I'll usually try and organise the scripts if they haven't come organised [referring to hard copies] according to the same order as the numbers appear on my screen ... So I can go through them systematically and not make any mistakes".*

'Alex', in common with a number of interviewees, talks about making notes on a sheet of paper:

*"I have a sheet of paper for each student ...when you are a new marker it is really scary and I found I was ... trying to read an assignment getting half way through and I couldn't remember what they had said at the beginning, or getting towards the end and couldn't remember what they had said in the middle. So for me it is important to make notes as I read through the paper".*

'Kim' is also a note maker, scribbling on a pad whilst reading the assessment artefact. When they finish reading it, it is then that they are able to complete the feedback sheet by returning to the comments made on their note pad. They approached construction of the feedback by only writing comments in an overall section, at the bottom as a single element rather than make individual comments. 'Kim' demonstrates the difficulty posed by areas that require evaluation according to the marking and grading criteria, for



instance separating style and structure for individual commentary, perhaps viewing these things as inseparable.

On another note, “Chris” commences with checking the assignment brief given to the students, so that they are clear about what the completed assessment artefact should comprise.

‘Gerri’ describes their own approach:

*“I assign them all a mark as I go through, note it on a piece of paper or write it on the script, we are not supposed to put comments on the actual script”.*

Other interviewees discuss the practicalities of undertaking the task, which to those individuals hold some level of importance. For ‘Danni’ one such task was related to the correct presentation sources of used in the assessment artefact:

*“...I have the word document open. So I tend to check the references first. **[to be able to action this]** I have two documents open and I go through the text and when I come to a reference I highlight it in the second document”.*

This interviewee works at an institution where electronic marking and grading has been instituted, making the checking of a reference list a more complex procedure than when the artefacts were submitted on paper.

‘Alex’ confirms that this is a legitimate task, by articulating their process:

*“... I am also looking for underlying technical stuff, are they using the Harvard **[referencing]** system correctly*

These interviewees have instigated a set of mechanisms to facilitate them evaluating written assessment artefacts, including checking what they view as vital components. They have been able to articulate the processes they implement in order to be in a position to begin evaluating a single or an entire batch of assessment artefacts.

#### 4.3.1.c HOW DO YOU MARK OR GRADE A STUDENT PAPER (ASSESSMENT ARTEFACT)?

This sub theme examines interviewees' approaches to marking and grading a written assessment artefact. It reveals the nuances in individual lecturer practice when faced with the task of evaluating the assessment artefact.

A number of interviewees discussed what they looked for in an assessment artefact. 'Alex' has an approach where they actively want to award marks or grades:

*"... going through looking for I'm looking for where I can give marks for knowledge understanding and application ... the people get the higher marks have read around the subject and often they are bringing in new knowledge".*

Other interviewees discussed observing the style in which the assessment artefact had been written. 'Alex' sums this up well:

*"Are they are they writing in a manner that they are getting is it written in easily understandable English. You see is it easy to read is it easy to understand. Are they making it easy for me as a marker to give them marks".*

Meanwhile 'Gerri' has a different impression of what reaction the evaluation and feedback will receive:

*"We are allowed if we feel we can put a number and put something on the comment sheet but I tend not to do that either I am not convinced students read these comments we make anyway".*

The rationale 'Gerri' offers for this perspective is:

*"...they are a one off they are not going to be doing this again ...".*

It would seem that 'Gerri' views the process of drafting comments to provide feedback of their evaluation as an unnecessary task. They conclude that feedback at a summative assessment point to be of minimal value to students.

These next interviewees discuss some of the components that they look for when evaluating during marking and grading, related to the structural aspects first, then content.

'Chris' comments:

*"... things like referencing or grammar or paragraphing you know more around the organisation and structural aspects of the work. As you get into the assignment ... things about the content and such like but at the start it is about this assignment"*.

'Kim' reflects on the process of evaluating an assessment artefact which:

*"... I found very easy to mark ... it was from my field of practice so I very clearly had the facts at my fingertips cos I think that is often the struggle ... it was a joy to read because all the things I might have put in he did ... he integrated the person to the case study ... he gives a clear explanation ... also you know he critiqued the evidence really well ... looked at the relevant trials ... he had got it down very succinctly and he had written really well and I remember the piece being well structured"*.

Interestingly 'Kim' introduces the perspective of being able to complete the task of marking and grading more easily because of their familiarity with the topic area. This passage also incorporates all the components that previous interviewees have highlighted as important when evaluating an assessment artefact.

'Eddi' completes a similar process to other interviewees, but has assessment artefacts which they have printed, they then work using this paper copy:

*"...I will annotate and make notes, I might ring words if there is a grammatical error or something that is badly phrased I might ring it or spelling error or if there is something that's there's an issue with it then I might make a note down the side [margin] just to remind me when I get to the end and I come to mark it on the rubric I have then got my key points to act as an aide memoire"*.

Interviewees then go on to discuss the issue of interacting with the assessment artefact, Chris says:

*“I often find the first one is a nightmare [to mark] and then you get a bit more into the swing of it”.*

Towards the end of the marking and grading process ‘Gerri’ discloses how they will review the initial or provisional score awarded by briefly interacting with the assessment artefact for a second time to satisfy themselves that they have been thorough and justify their decision. They then go on to explain their process:

*“... you know I think there is some norm referencing with the group you know thinking ... if I am giving that a mark of eighty and this one here is getting a mark of forty first of all is that really the best script in that pile you know that's what I want to make sure ... that I am being fair”.*

#### 4.3.1.d USING THE MARKING AND GRADING SCHEME

Interviewees were asked to identify their approaches to using any provided marking and grading scheme or criteria in their evaluation of written assessment artefacts.

‘Billie’ suggests that reviewing the marking and grading rubric will increase the lecturer’s understanding:

*“so you get a sense of what a really good essay is and you can get a sense of what a medium essay is ... and you know what a failing essay usually is”.*

While ‘Terri’ indicates that in preparation for marking and grading that they will:

*“... have a look at the mark criteria looking at what we have asked the students to produce so I have got something to judge it against”*

To effectively make use of the marking and grading criteria ‘Mel’ says:

*“... I have read through it once [the assessment artefact]... I go back through it again and pick out, sort of allocate the scores for each line of the grid and add it up and divide it”.*

From this excerpt it is evident that 'Mel' like 'Gerri' proposes to read each written assessment artefact on two occasions to complete the marking and grading evaluation. Whilst this does in effect double their workload, it does illustrate the operationalisation of analytical criteria, by systematically applying the specified standards to the assessment artefact.

To have increased certainty in applying the marking and grading criteria in a confident way 'Lesli' reports how they needed to be familiar with both the assessment criteria and the assessment artefact. They then reviewed one against the other i.e. the criteria and artefact, by asking questions of the assessment artefact, relating to how well 'Lesli' thought they had achieved the requirements:

*"... if you start out at seventy and work a way back ... I think if you have got 70 as a first column that is the ultimate [70 is seen as the ultimate mark to achieve]. And then you start to think they haven't hit that but they are down here somewhere".*

In effect this interviewee reviewed the criteria by starting with the perceived highest mark or grade and only moving down to the next lowest section on the rubric if there was no evidence of achievement. This process would be continued until 'Lesli' considered they had reached the criteria which could accurately describe the artefact in the language of assessment.

'Billie's' perception adds a further dimension:

*"... essentially it is how you work out all that on the rubric, cos that rubric shapes what you do. But the rubric doesn't [always] fit perfectly well".*

Yet 'Chris' is able to suggest how the marking and grading criteria and rubrics support the crafting of written evaluation of the assessment artefact:

*"The marking criteria is great. I think it helps a lot. It helps us to make sure we comment on everything that needs to be commented on. So it is a bit of a prompt I suppose from the marker's point of view".*

Further to this 'Gerri' has an interesting perception on the impact of the criteria, when they do consult it:

*“... you might have a look at the criteria but I always find I am looking at the criteria quite negatively I am looking at the criteria to find ways to say ‘they didn't do that’, you know what I mean. I find it difficult to discriminate between sixty four and sixty six”.*

‘Gerri’s’ approach is not one mirrored by ‘Chris’ who, in comparison, is much more positive of the utility of the marking and grading criteria. Interestingly both of these lecturers are experienced in marking and grading with experience in HE far in excess of the two years (the description of a novice lecturer) (See 4.2 Interviewee Key).

‘Billie’ raises an issue with the use of the same criteria to mark or grade different types of assessment artefacts:

*“... we have to use the same rubric so we use the exam rubric for the protocols because they are submitted under exam conditions but when I have used those to mark a dissertation with they don't work because they are not designed for an assignment [**written assessment artefact**]”.*

‘Terri’ states the importance of familiarity with:

*“... whatever feedback mechanism you have got available to you so whether that be a rubric or whether you can mark on script and then how you are going to go about doing that ...”.*

One interviewee makes use of normative referencing, ‘Gerri’ explains that whilst being aware of what they refer to as an “*objective marking standard*”:

*“... I like to get a feel of where things lie relative to one another”.*

They then discuss their approach which appears to be relative to their involvement in defining and developing the criteria, perhaps resulting in their disengagement:

*“... we have marking criteria ... I have not been involved in drawing them up but I have to say I tend never to look at those you know and never scrutinise them in advance of actually marking the scripts”.*

Whilst ‘Lesli’ considers that there are important features which make the difference between effective and ineffective assessment rubrics for marking and grading:

*“... you look at the wording in it and this is why marking criteria has got to be very carefully worded”.*

An interviewee at one institution ‘Mel’ discusses how they have individually designed marking and grading criteria:

*“... No it is per assessment they’re similar it is always that there is about five or six **[sections]** and the final one is always about spelling, grammar, referencing, layout but the others are different depending on the assignment ...”.*

‘Mel’ when reflecting on a specific evaluation task reveals, that deeming the assessment artefact as being unsuccessful was not a difficult task:

*“... It was easy to fail because it clearly says on the marking criteria that you will get less than forty points ... for that section if you do not do this ...”.*

In one way or another, these interviewees are exploring the constituents of the criteria or rubric for completing a marking and grading task. They identify how fit for purpose or useful they may be in underpinning decisions made by those completing an evaluation of an assessment artefact

#### 4.3.1.e ENVIRONMENT

Interviewees explored how the environment helped or hindered their ability to complete the marking and grading task required of them. Excerpts from five interviewees have a fairly common theme, with many of the interviewees opting to complete their assessment tasks away from a shared office environment.

‘Danni’ is the first interviewee to indicate that a shared office is a problem when marking and grading needed to be completed:

*“... Cos I talk to myself I think and I get up and I am a fidgetter”.*

When in an office being shared with anything from one up to nine or more colleagues, having someone constantly getting up and down, moving about, shuffling papers, rearranging pens etc., ‘Danni’ felt that this would be a severe distraction to others.

'Mel' concurs with this perspective, they plan their workload when they have marking and grading coming in so they:

*"... can do it at home with less distractions cos it is impossible to mark in an office with eight people".*

They go on to give their rationale for this, being in an environment where:

*"... people are constantly talking to you and stuff".*

This leads to an environment which is not conducive to the levels of concentration which are required to complete the marking and grading task, needing time away from the work place to do this.

Another interviewee 'Kim' finds that as well as needing:

*"... a quiet place so I can concentrate".*

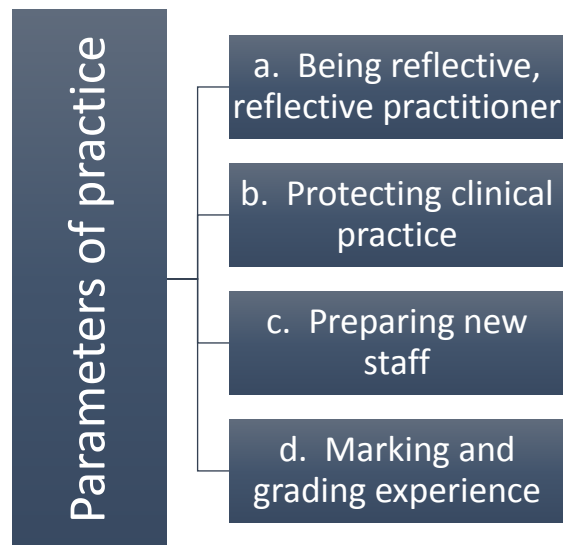
They are also aware of what other aspects facilitate their levels of attention and concentration along with being away from a shared office environment:

*"... I usually do it in the morning cos I know I'll concentrate best then".*

These interviewees demonstrate that it is important to consider the working environments in which those completing marking and grading find themselves. From completing these interviews it appears in the majority of cases the propensity to work in a shared office is not conducive to completing the task. However, two interviewees, whilst they had a single person office felt the potential disruptions meant that somewhere away from the formal office atmosphere would be their environment of choice.



## 4.3.2 PARAMETERS OF PRACTICE



**Figure 7 CI Major Theme**

This second major theme identifies a number of elements which the practice of marking and grading may be subject to. The lecturer completing the marking and grading activity may be influenced by one of the suggested conditions categorised by the sub themes.

### 4.3.2.a BEING REFLECTIVE, REFLECTIVE PRACTITIONER

The concept of being reflective was identified by three of the interviewees as being important in the evaluation of written assessment artefacts, with each of them taking a slightly different perspective. For instance ‘Billie’ made a distinction between being a reflective practitioner in their previous existence as a health care professional, and then as an educator. ‘Billie’ was very careful to ensure I was aware of the difference:

*“... as a reflective practitioner as an educator. Not as whatever it was we were all before we came here. If that makes sense. I think that is really crucial ...”.*

This interviewee was talking about the concept of reflection in relation to marking and grading, and how it is useful to review a paper even following arriving at a final mark or grade. However, ‘Billie’ said that at times:

*“you go back to it again and you go ‘Need to change that’. I think that is a really important fail safe and that is a really important component of being a reflective practitioner...”*

‘Billie’ considered that there could be a number of things that might have an adverse impact on their ability to make the assessment process one which was of benefit to the student. Reflection therefore was seen in this context as:

*“... really key. Cos if you are distracted and you perhaps [are] less focused on making sure people really grow out of this”.*

This reflective activity had the power to re-awaken memories for ‘Chris’ and their own experiences of being a student:

*“I think how I felt when I got results and I'd think Oh god I think it was better than that”.*

Conversely Chris recognises that students are able to clearly identify what is required of them due to access to clear criteria and the availability of formative feedback related to the assessment artefact. Despite this position ‘Chris’ also explored the ramifications for the student of receiving an inadequate mark or grade to warrant a pass in the assessment task. However ‘Chris’ also added an interesting dimension to their exploration of reflection that of remembering individual students who had perhaps been unsuccessful and the feelings this evoked:

*“...and it is about their chosen career and it is about their working life. It is about who they are and what they are so it is really important to them so. You know that is why I think I can't look into their faces ...”.*

Finally ‘Chris’ also went on to give further evidence of their in-depth understanding of the student’s position. Again taking a reflective stance enabled ‘Chris’ to empathise with the amount of work the student had possibly completed, yet to not achieve the success they desired:

*“So it must be devastating cos to you it takes a lot of effort to write pieces of work and so you know it is not a trivial thing is it for them”.*

Another interviewee, ‘Alex’ broached the concept of reflection from the angle of experience and its impact upon marking and grading practice. Their

perspective related to participating in this research and they saw this as a valuable opportunity to closely examine a taken for granted activity:

*“..because I think when you do become a marker with experience I think it is very easy to sort of subconsciously go along not thinking about what you are doing and just redo the same things over and over again. And I think having the opportunity to reflect on my own systems is useful and then you think about rightly or wrongly ... sometimes you do think is it right that I do this or is it wrong?”.*

The subject of reflection prompted interviewees to explore their experiences of marking and grading. Interviewees took the opportunity to take an empathetic stance, putting themselves in the shoes of the student, imagining their feelings. All interviewees, apart from ‘Alex’ imagined the feelings only of students who had been unsuccessful or less successful than they perhaps had wanted, not on what achievement beyond expectations would feel like. ‘Alex’ was the only interviewee who explicitly reflected on their process of marking and grading without relating this to a particular evaluation event.

#### 4.3.2.b PROTECTING CLINICAL PRACTICE

The subject of ensuring students had the potential to become safe practitioners, thereby protecting clinical practice and protecting patients, was identified by four of the interviewees. ‘Lesli’ cites the impetus for protecting practice:

*“You have a moral obligation and a professional obligation. You also have this thing called the code of conduct that guides us as to what we do”.*

The Code of Conduct being discussed is published by the Nursing and Midwifery Council (NMC) (2015) and guides all professional nurses and midwives practicing in the United Kingdom. Every practising Nurse, Midwife or Health Visitor in the UK has to be registered with the professional body, the NMC. In spite of ‘Lesli’ no longer being in clinical practice, there is a requisite to maintain a live registration with the NMC, meaning The Code continues to be a living document for this individual. Hence it becomes evident that academics who remain registered with this professional body

review the assessment artefact as a subject expert from two perspectives, their role within the institution along with a responsibility for protecting the integrity of their profession.

'Mel' highlights the possibility that assessment artefacts which present clinical issues are missed. As a lecturer recently arrived from practice, the potential of clinical errors and their detection remain at the forefront of their mind:

*"Some people have got a fantastic academic style and then you read it and think that's a real patient safety issue or that is a safeguarding issue that you've not picked up on or that is actually abuse or that is wrong. That's you know the clinical stuff and people miss that. Because the academic side of it [it is well written]"*.

This position is confirmed in comments made by 'Lesli':

*"...but the thing is we've got to protect patients and our safeguarding. And some of the horror stories you get and you think I don't want them near patients"*.

These two lecturers appear to suggest that students may write about clinical issues, but could give clues as to their inability to transfer what they write into the practice environment effectively. Or conversely what they write may be illustrative of a poor approach to patient care in practice. The dual role here taken by academics with a professional registration seems to bring them into conflict when assessing an artefact which could meet the requirements of the institution yet be contrary to the requirements for safe healthcare practice.

'Terri' takes this one stage further by equating the achievement of the assessment artefact learning outcomes as evidence of success in practice:

*"...actually has she answered the question has she achieved the learning outcomes will she be able to do these things as a result of doing that assignment. And that I think was a bit of my kind of pass/fail line for that student so it wasn't .... what points did she miss ... but has she answered the question with enough depth to be a decent clinician in that particular environment..."*.

This presents evidence of the dilemma that can be faced by lecturers with a dual role as both clinician and academic. 'Terri' considered that the assessment artefact they evaluated provided evidence that the author possessed enough clinical knowledge, but was nevertheless unsuccessful at the assessment event, because of a paucity in academic prowess.

#### 4.3.2.c PREPARING NEW STAFF

The role and opportunity to prepare new staff for their marking responsibilities was discussed by three interviewees. 'Harri' gives an explanation of a recent experience of trying to teach a new member of staff. They described how both 'Harri' and the new staff member would share what they had found; they would separate for a time to mark or grade a paper and then come back together:

*"...[to] share..... , well if we've not been sort of on a par I've been able to explain to them why I've gone down the way I have gone and the rubric does help me to do that and I think you do need something like that to, particularly if you're teaching someone".*

A reasonable approach from 'Harri's' perspective is to encourage the new member of staff to utilise their previous knowledge and experience:

*"Cos they're all practitioners, they've all got experience in their own right..."*

'Harri' seemed to suggest that it was appropriate for the new lecturer to be able to apply their previous clinical practice knowledge and experience to marking and grading an academic assessment artefact. An introduction to marking and grading is thus not provided, as is seen in the extract below:

*"... See what you make of it and then you tend to do it that way and then we'll compare notes and, rather than sitting down with somebody and saying this is Grade Centre this is what you are looking at, this is what you're looking for. Erm it tends to be more a case of this is the learning outcomes, off you go and it's quite, erm I don't know ...".*

Another interviewee 'Mel' describes their experience of being introduced to marking and grading with being provided with limited if any support to develop effective marking and grading practices:

*“Well the first few I marked I went to my line manager and got him to double mark them and check that because I had never marked before and you don't get taught how to mark and there is not any support as such”.*

The experience of another novice marker mirrors that of the above interview. ‘Jamie’ relays their induction into this and other aspects of their role:

*“**[What support did you get?]** None. The first bit of marking I had took over the brand new module the module leader left. And I had 120 exams to mark. So I didn't teach on the module and so what we did I split it up and across the whole team”.*

This process is implemented to support those new to marking and grading. A rationale for this is explained by ‘Harri’ in two ways, the first being a time resource issue, though the second point put forward is somewhat unexpected:

*“I think it's a time issue you don't have the time to sit down with somebody to go through the thing in that detail and course if you don't feel overly confident about your own marking you don't necessarily want to relay all your thoughts on to them”.*

‘Harri’ appears to highlight that a lack of confidence in their own marking and grading approach is, for them, one of the issues which prevents a closer collaboration between seemingly experienced academics and those new to the role.

Whilst another interviewee expressed that they needed to make a concerted effort to seek support for learning to mark and grade due to the position they occupied as a new marker, early in their academic career. ‘Mel’ conveys this in the situation they found themselves in:

*“...and so then you'll say right all me marking has come in now. Will someone buddy up with me and go through it with me and then no one can and it literally feels like you are just hit the ground running ...”*

A further observation by ‘Harri’ concerns life beyond an introduction to marking and grading, to that of continued professional development and

improvements in practice. They observe a lack of available resources on the actual practice of marking and grading:

*“...there isn't, there don't seem to, to be the, to be resources. There's lots of workshops and things are going on, on other things but not in relation to marking when the marking is the bit that has a real big impact on their **[students]** future”.*

The common theme expressed by interviewees was a lack of preparation for their new roles and a lack of appreciation of the anxiety this caused. It appears those marking and grading with a greater length of experience did not establish that there was to be a significant learning curve in making the transition from clinical health care practitioner to academic. An assumption was made in a few cases, that practitioners would be able to use their previous clinical expertise in transferring this to the evaluation of written assessment artefacts.

#### 4.3.2.d MARKING AND GRADING EXPERIENCE

The majority of interviewees disclosed their length of marking experience (Interviewee Key 1: 4.2). For some interviewees it was this feeling of experience which facilitated them in reviewing the submitted artefact and then coming to a judgement in relation to awarding a mark or grade.

‘Alex’ also notes that there could be far reaching consequences beyond this assessment period:

*“I am an experienced marker and I think there are traits you can see in assignments that you are thinking if this student doesn't get it now this is going to be a theme... **[an]** issue that is going to rumble on in the second year and possibly end up with them failing the third year and not being able to go on to the jobs market or **[to have to]** leave the course”.*

From ‘Alex’s’ perspective, with experience comes the risk of viewing marking as a repetitive act, one which required little conscious activity and attention:

*“I think it is very easy to sort of subconsciously go along not thinking about what you are doing”.*

'Alex' seems to be describing being unconsciously competent or not engaging with the assessment artefact in a mindful way.

'Billie' then goes on to illuminate this further by explaining or describing the process of moving from limited understanding to a position of proficiency in marking and grading as a rite of passage:

*"... so you get a feel I think over time of doing things and I suppose it is that bit of novice to expert journey".*

'Billie's' journey from novice to expert had taken in excess of 5 years and their experience was gained at two universities. They identify the different support in place for marking and grading written assessment artefacts by comparing the available marking rubrics:

*"I think they are good they are good [At their current institution]. Whereas the others were so broad and lose you could have just said anything really. There was nothing to help you keep consistent."*

When trying to unpack further the concept of experience and how this is portrayed in the lecturer knowing that the marking and grading is right, 'Billie' asserts that this is simply through experience. They consider that being exposed to the right experiences as being important and this comes in the form of communication with others:

*"...talking to other people about where they think this goes and what and why. And that is how you develop your experience and it is kind of its reflective, its experiential, you kind of engrain some of it tacitly don't you and it is that. It is a mixture. It is an absolute mixture".*

This is supported by 'Chris' who has been in education for 21 years and links their length of experience with being able to complete the task of marking and grading efficiently:

*"I very rarely read it again. I know a lot of people read it once and then read it a second time to mark. Very rarely. I don't know whether that's due to having so much experience or maybe I am confident in what I do.....".*

'Chris' explains how, when they undertake marking and grading, they can do this by only reading the student assessment artefact once. In that single



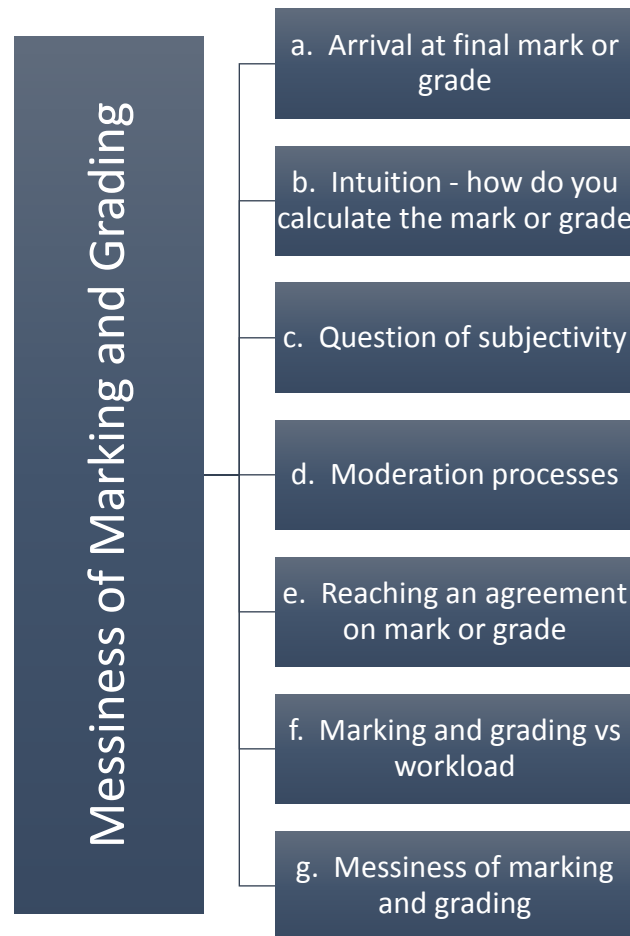
reading they are able to provide a mark or grade and assemble feedback for the student. This is not the case for a number of their colleagues, especially less experienced markers.

'Mel' a very new academic with less than 1 years' experience talked about having the foresight to plan ahead for marking periods, expressing the concern that it would take them longer than their more experienced colleagues. They talked about trying to weave the marking in and amongst other commitments such as teaching, student support, meetings and delivering mentor updates in clinical practice. Planning ahead is only effective if you are aware of what the demands on your time are to be. The anxiety that was evoked when 'Mel' was faced with things out of their control is evident in the following passage:

*“But then they produced this grid with all the marking on of who was going to mark what and when and there was loads of extra stuff on there that I had never been told I was going to mark and then extra assignments kept appearing that had not got in my diary..... Cos I didn't know it was coming”.*

This sub theme has explored the thoughts and feelings of interviewees with a range of marking and grading experience. The majority of interviewees who were not identified as novices in relation to marking and grading felt confident in their ability to evaluate assessment artefacts. However, not all talked of this confidence transferring across to support the development of newer members of their academic teams. Without exception the interviewees considered experience to be the key to effective marking and grading practice.

### 4.3.3 MESSINESS OF MARKING AND GRADING



**Figure 8 CI Major Themes**

This major theme explores what are viewed here as contextual practices which impact upon marking and grading. The sub themes attempt to explain and categorise descriptions from interviewees of aspects of their marking and grading practise.

#### 4.3.3.a ARRIVING AT FINAL MARK OR GRADE

Once marking and grading of an individual script is complete, exploring how the final mark or grade is calculated provides particular insights into the complexity of arriving at an evaluation for an assessment artefact. One interviewee, 'Mel' discussed how they had honed their skills in providing an end grade by comparing their marks for three assessment artefacts with those of their manager. They were both satisfied when the end marks were

only a few points different. However, 'Mel' also disclosed a comment made to them about awarding marks:

*"...I think my manager said to me something on the lines of I never give, if I give more than 60% I am in a good mood and if I give more than 70 I must be drunk".*

This extract portrays the use of normative referencing, rather than criterion referencing. 'Mel's' manager potentially has a glass ceiling for the awarding of marks or grades and it is unusual for an assessment artefact to transcend this.

After marking and comparing the marks with their manager 'Mel' gained in confidence, feeling that they were on the:

*"...right track ...".*

A patchwork approach to marking and grading became evident when 'Mel' talked about the difficulty of being able to consistently evaluate the quality of the work in front of them:

*"...I found I might mark one and give it one score and then mark another and then go back to that one and think oh actually that one I thought that were really good but now I have seen these few that one is not so good and bring it down...".*

A similar process to this is acknowledged by another interviewee, 'Lesli' who describes how they arrive at the final marks or grades for a batch of assessment artefacts:

*"... when I am marking what I tend to do is I tend to mark four or five and then if I come across a really good one I will go back and look at the others and I'll grade it against that really good one".*

Whilst another interviewee 'Harri' discusses the difficulty with arriving at a final grade when:

*"... there seems to be some content there and, but it's very difficult to pinpoint whether the student knows little about or whether they're just describing what's going on ...".*

Another aspect of this is echoed by 'Billie' who explores the dilemma of making an assessment of an artefact which is written well:

*"... people write really well express themselves well but actually have no substance to what they are writing, [a] really nicely written piece of work which doesn't answer the question I think actually is a harder one to manage ..."*

'Billie' also explores what they refer to as the "veneer of goodness" and how to deal with this facet when completing marking and grading:

*"... [A] really nicely written piece of work, which doesn't answer the question I think ... actually is a harder one to manage ... you have to get past that veneer of goodness ..."*

A further point worthy of interest was 'Mel's' perspective on the diversity of marks awarded by colleagues when marks were always significantly different and this appeared to be related to the way they used the assessment criteria:

*"Whereas somebody else gives nineties out because she marks down and he marks up"*

This sub theme explores how interviewees arrived at the final mark or grade. A greater propensity for comparison has been identified, either with the outcomes from more experienced academics for novice lecturers, or between artefacts in a given batch for evaluation. This illustrates the process of normative referencing.

#### 4.3.3.b INTUITION – HOW TO CALCULATE MARK OR GRADE

When interviewees discussed how they achieved the final mark or grade to be awarded for assessment artefacts, in different ways they explored how they often had a feeling about what the mark would be. Eight of the interviewees in their own way expressed what I have called intuition, in being able to know what the mark or grade should be and would have a level of concern or curiosity if this was not achieved. 'Alex' discussed how on reaching the end of evaluating a written assessment artefact that:

*"So generally I find that it is in the correct ball park but sometimes it isn't, then I think why isn't it?"*

Interviewees in this case then discuss how they would review the work

and/or their written comments about the assessment artefact to support the mark or grade awarded. 'Alex' identifies how it is not an impossible situation to award something different to what they had in mind at the outset:

*"I have written one thing and given a mark which is completely different so I think sometimes there are just simple errors but sometimes you look back and think maybe I have been a bit tight there, maybe I have gone to the lower end of that scale and maybe I could go to the middle of that scale or something".*

To deal with those differences or incongruities it necessitated lecturers undoing their original mark or grade, 'Alex' puts this very simply:

*"So I do change the rubric when I compare the mark to the one in my head".*

One interviewee, 'Chris' is much more revealing when talking about how their process of insight works, from initially beginning to read an assessment artefact and thinking that they know the value in terms of a mark or grade awarded:

*"... well as you are reading through you, you tend to start to think, oh this will come out around about. You get a feel for it..."*

When this 'feel' occurs is dependent upon the length of the assessment artefact, so for a short piece of work 'Chris' asserts that it can be early on:

*"... into that piece of work you can start to get a feel for how the student can write. How they are putting things together..."*

It also appears that lecturers are able to identify where a student and their assessment artefact may fit, and this is discussed in relation to the work itself and how it is put together. In fact 'Alex' echoes sentiments of another interviewee by saying:

*"... there is a sense of fitting [**a particular mark or grade**] because if someone writes in a very coherent way and they have expressed themselves well..."*

This feel extends further than simply from coherence, but to include the way students are able to explain the subject matter, but doing more than using description, this facilitates an insight into the level of understanding of the

artefact topic. In isolation this comment appears to suggest the ability to write well and to write coherently as a good indication of identifying where the artefact may fit in terms of marking and grading the piece of work. The institution where this individual is employed uses online marking criteria where coherence, along with attention to structure make up a small proportion of the available mark or grade that can be awarded. Interestingly 'Chris' considers that it could be possible to gain a fairly accurate picture of achievement in the early part of evaluating a written assessment artefact, to be able to make a note of that estimation and to be in the same ball park at the end:

*"... I suppose in some ways you could say to yourself well write a number down part way through and then go through **[to the end]** and do the marking criteria and see what comes out. The chances are it is probably very similar to what you think".*

There are those who could think that this occurs by chance or is a self-fulfilling prophecy. This was a sentiment expressed by more than one interviewee; 'Harri' for instance expresses this in terms of being able to sift through the different levels of achievement:

*"I suppose you get into the habit when you've had years of practice, that you know what's a good paper and what's not such a good paper".*

The role of intuition in constructing the mark or grade appears to be more prominent for those academics with more experience, however is not confined to that group. Interviewees demonstrated that intuition alone did not form the whole basis for the award of the final mark or grade, but was definitely implicated. Interviewees were more likely to return to the assessment artefact, there appeared to be an incongruity between their 'feel' for the final outcome and that which was achieved by completing the marking and grading rubric. Whilst interviewees described their process or actions to be guided by 'feel', I have purposely renamed this as intuition. Interviewees seemed to be describing their knowledge or actions as something innate or instinctive with limited ability to provide a clear explanation or rationale.

#### 4.3.3.c QUESTION OF SUBJECTIVITY

This subject was raised by five interviewees who all, in one way or another, discuss the internal deliberations that occur regarding a mark or grade they are to award to an assessment artefact. The interviewees spend time considering whether the mark or grade is reflective of the assessment artefact or if they have judged it too leniently or indeed too harshly. They each pitch what they call their subjective judgement against taking a more objective stance, appearing to favour use of an impartial, dispassionate approach instead of preferencing the individual and particular. This is well articulated by 'Alex' who states:

*"... sometimes I disagree with myself and then what I have to do then is debate whether my subjective thoughts were more accurate than my objective, than what the objective device has given me and that is useful ..."*

They then go on to explore this further by trying to make a case for why subjectivity amongst those marking and grading exists:

*"... everybody thinks differently ... everybody has an idea of what they want ... I think the rubric tries to bring objectivity to the process ... one standard set of criteria ... doesn't stop me as an individual being different to another person's individual but at least you have got some sort of common ground which is clearly explicit ..."*

'Terri' highlights how, through the uncertainty they try to aim for as greater precision as possible when constructing the mark or grade:

*"... I think that I can say that I have tried to make it as accurate as possible through checking marking initially with colleagues and through kind of collating as much as I can in terms of feedback to check it against but I don't know what accurate is ... I mean it is not like an MCQ [**multiple choice questions**] where you can accurately work to a single percentage but I would like to say it is in the right grade band ..."*

Finally 'Mel' highlights that it was often difficult to mark or grade assessment artefacts because of not knowing whether the knowledge they were anticipating should be present was expected, they wondered:

*“So am I asking too much ... is that too high a level for level four **[the level that was being evaluated]**”.*

Having this approach potentially effects the evaluation of the written assessment artefacts which 'Mel' has been given responsibility for completing. Being able to pitch the expected merits at the correct level ensures a more appropriate estimation of the value of the submitted artefact; therefore impacting on marking and grading quality as well as potential objectivity.

This sub theme has illustrated the various perceptions of subjectivity related to providing an evaluative mark or grade for an assessment artefact by the interviewees.

#### 4.3.3.d MODERATION PROCESSES

This topic was raised by seven interviewees and was seen as an important aspect of evaluating the written assessment artefact. A number of interviewees expressed the need to use moderation processes to assist them in confirming the mark or grade to be awarded. All interviewees who raised this issue identified moderation as a process of getting another, possibly more experienced lecturer to review the mark or grade awarded, this process was being called second consideration. The word moderation in itself is related to restraint, control or temperance. Those interviewed expressed a requirement to ensure that they were being fair to students and that they were being equitable. Most institutions have a set standard for moderation of artefacts submitted for assessment and it is not uncommon for this to be a small percentage of the total assessment cohort.

'Billie' for instance picks up on this point:

*“No it goes on the sort of the fails and the ten percent rule really”.*



Although they also identify that moderation is not a fool proof process, with individual interpretations still forming a greater part of providing the mark or grade for the assessment artefact. 'Billie' highlights this as:

*"...how long is a piece of string? ... because it is still my judgement at the end of the day and that is when I rely on my colleague to review and I review that lecturers and then you will come up with a you know [mark or grade]. We very rarely swap the grades but we do occasionally".*

Interviewees felt most benefit from moderation when they were unsure or uncomfortable with a mark or grade they had awarded, 'Alex' highlights this in saying:

*"It was a difficult call and I did get someone else and we do second consideration but it was one I asked for second consideration on".*

Whereas novice markers and graders gained some level of reassurance that their attempts at evaluating an assessment artefact would be confirmed as correct via a moderation process, as Jamie identifies:

*"I just follow my process, at least then I know that my marks ... obviously being fairly new it does get moderated ... so far don't think anyone has really changed anything".*

In fact interviewees disclosed that they would often instigate the moderation process when they had assessment artefacts which lay on the borderline of either pass or fail or indeed a border for a good classification. 'Billie' points out:

*"... If people are on the boundary line I will also sometimes go have I been really strict this time round or have I been a little too easy going. So I will ask for my boundary people to be checked".*

Interviewees use this process for support when they were unsure and had some level of anxiety about the mark or grade their evaluation had generated. Such anxiety arose due to a mismatch between the final mark or grade awarded to the assessment artefact and what the interviewee had considered would be a fair evaluation of the work. 'Alex' highlights how they use the second consideration process to assist with coming to a better

decision, than they could make individually, as to the worth of an assessment artefact.

Instigating second consideration does not always mean that there will be agreement on the mark or grade awarded for an assessment artefact as 'Chris' points out:

*"... in fact I was considered by the moderator to have been fairly generous and ... it was a supplementary I was looking for the work to have improved from the first submission but it had actually [got worse]..."*

'Chris' then goes on to explain why the over generosity had occurred:

*"Yes like I said the moderator said I had been a little over generous but I suppose I was just trying to find something for the student"*

'Billie' made an interesting observation about the constitution of moderation teams:

*"... you would have to moderate marking teams so we worked out a system to try and get consistency because you would always have the really high marker and the really low markers and you would have to pair them off so they could battle it out..."*

In reality once marking teams are established (for the duration of the assessment), consistency is possibly not achieved for all submitted assessment artefacts as moderation, in the form of second consideration, is for a defined sample only. To do any other would necessitate all assessment artefacts being marked or graded by both members of the moderation team, thus doubling the workload.

Different concerns were raised by those with less experience in marking and grading, 'Mel' for instance describes feelings of inadequacy in the role of second considerer. This was particularly evident when participating in a moderation meeting where a number of lecturers involved in completing marking and grading for a particular assessment artefact will hold a discussion. The observations of this interviewee are interesting:

*“I have noticed I moderated for the first time the other week. And I didn't disagree with anyone's scores because I was under the impression that you don't do that. And who am I to do that cos I am new anyway. But there were some saying oh you are a bit harsh, they're only at level four. Or ooh I can see a big glaring thing there you have not picked [th]em up on. Or you know I didn't think that was as good as you did and it's just such a personal opinion”.*

It seems as a novice marker or grader ‘Mel’ did not feel encouraged to question, let alone challenge marks or grades awarded by those with greater experience.

Whilst ‘Alex’ discussed how marking and grading teams seem to be established in their department, but for this current round they had been paired with somebody that they had not completed any moderation with before. ‘Alex’ appeared to be excited rather than apprehensive at the prospect:

*“but this year I am marking with someone I have never marked with before who is an extremely objective person so for this module I marked half the assignments and this other person marked the other half and then we came together and we second considered each to sample that was quite useful cos it was somebody different”.*

Moderation is viewed by the interviewees as a process by which they can gain confidence and surety in the marks or grades awarded to assessment artefacts. The fact that not all artefacts are subjected to this double checking process to increase objectivity does not seem to be highlighted as an issue to more than one or two of the interviewees.

#### 4.3.3.e REACHING AN AGREEMENT ON MARK OR GRADE

In relation to reaching an agreement on the mark or grade to be awarded, 5 of the interviewees discussed this topic. They deliberated on a number of issues which could impact on how agreement was reached on what the assessment artefact would be awarded. These considerations appear to come into the following areas: issues in applying the whole available mark or grade, academic level of student, knowledge of the assessor/ ‘expertness’ of

the assessor, size of the variance between the mark or grade of the two lecturers, and individual interpretation of the learning outcomes.

'Alex' highlights the difference in the approach of two lecturers to a first year assessment artefact, identifying that what is required is different to students further on in their academic programme. This interviewee supports the notion that it is:

*"...much more you have got to accept that it is quite knowledge based at that stage and this person actually did demonstrate more knowledge than I had given them credit for [initially]"*.

This was 'Alex' reflecting on a disagreement with the moderator over an assessment artefact, when they had initially awarded a mark in the 40% band. However, the moderator wanted to change this to a mark in the 50% band to take account for the descriptive knowledge displayed by the student.

Understanding what is required from the assessment artefact, that is knowledge and level of expertise of the lecturer completing the marking and grading has an impact upon the evaluation of the item. When two lecturers come together to agree what should be awarded through the second consideration or moderation process, the level of expertise will come into view. One lecturer is more likely to possess greater assessment literacy, knowledge base and understanding of the marking or grading criteria, creating an unequal footing in expectations of what warrants a pass mark, 'Kim' highlights:

*"... we were both looking for very different things. And so um so I have a lot more problem there with applied path physiology and pharmacology."*

This perception exists, even in the face of explicit assessment criteria in the guise of marking and grading rubrics, which should be applied consistently by all lecturers within the department.

What became evident in talking with interviewees was the level investment in the module related to the level of ownership, with which comes a heightened understanding of what is required from the assessment artefact. This

provides those lecturers with a different vantage point from which to view achievement, either requiring more from the assessment artefact, with limited level or stage discrimination, or to subdue those with less expertise.

The use of vantage point is evident in relation to a different aspect of reviewing an assessment artefact; 'Gerri' discloses an event when second consideration was used:

*"...and then we meet to discuss ... not discuss the ones where we agree on ... but when we have a big disparity ... there was one I had given something like 80 and my pharmacology colleague gave it about a 40. I thought wow that's a big one I am not usually that far out ... but I have not understood something that they got wrong regarding the actual pharmacology, so I had just to say OK that's fine I wasn't getting that ..."*

The different perspectives of those involved in assessing an artefact lead to potential variance between marks or grades. 'Alex' highlights how the size of the variance influences the response of the lecturers involved in evaluating the artefact:

*"... when you second consider, if you come within three or four marks of each other I think that's a result. If you are in five or seven marks five seven eight marks then you'd need to ... just have a little bit of a think, if it is wider than that then you have got a problem"*

However what does become clear is that 'Alex' is discussing a process of double marking of an assessment artefact, which is completed blind, meaning that neither of the staff are aware of what mark or grade (and feedback) the other has awarded until the process has been completed.

Whereas moderation, by second consideration involves a slightly different process, in which the second lecturer reviews the grade and comments awarded by the first marker and then either agrees or disagrees with the original evaluation. Some evidence of the speediness and superficiality that the process of moderation by second consideration can take, is illustrated by 'Danni', who asserts that this is something that occurs on a regular basis:

*“... but XXXX second marked it for me, she cast her eye over it and came up with the same mark ...”.*

On initial review this process implies that a more accurate evaluation of the assessment artefact will be completed because of the measures implemented.

The process of reaching an agreement, causes anxiety to lectures just as does the moderation process, with those who view themselves as novice or with less experience finding difficulty with identifying why a variance might exist, especially when there is perceived expertise, as ‘

’ recounts their experience:

*“... I didn't disagree with anyone's scores ... cos I am new anyway”.*

The idealistic interpretations of how easy it is to reach an agreement as to the final mark or grade to be awarded to an assessment artefact are under represented by the interviewees in this study. What seems to be evident is that where there is perceived increased levels of experience or knowledge, those who take on the role of novice are less likely to disagree, issues of anxiety and power come to the fore. This sub theme contains more evidence that the process on which lecturer's rely to provide assurance is potentially flawed.

#### 4.3.3.f MARKING/GRADING VS WORKLOAD

Interviewees were concerned about the amount of time available in which to read through a written assessment artefact, then to provide a mark or grade and construct meaningful feedback. Comments on this topic are reported here from 9 interviewees.

‘Danni’ reports on how time limitations in the guise of shorter deadlines and shorter turnaround times, has necessitated a change in their practice:

*“... I used to check them all [references], I get a sense if someone does it really well or they're making lots of mistakes then I might periodically check one, I used to check the whole lot and I still do at undergrad but at post grad I tend to check the first few then I might have a look at the reference list ... and see how they have generally*

*written them ... then I shove that down and make my comments and then I start marking. So I always do references first get that out the way ...”.*

They then talk of how long it takes to read and mark or grade the different types and lengths of assessment artefact:

*“... It will take me, for this final stage with a 3,000 word journal article and a 12,000 word paper, it takes me a whole day. **[Level 7 dissertation]** ...”.*

Whilst ‘Kim’ identifies how they manage their workload and continuing to maintain an appropriate standard:

*“... but I also have quite a limit of how many I can do in session ...”.*

‘Danni’ goes on to explore, for them, why this is necessary:

*“... cos I am just so slow ... when I know it has been handed in I block a day off to mark at home. And unless I am absolutely made to come in that is it ...”.*

They then go on to describe what sort of day that might be:

*“... the 3,000 worders I am marking, I can do about three in a day. It’s a long day cos I don’t say I don’t have breaks cos I do, and I sometimes I have to move away from it, I think let me go process that...”.*

Whilst ‘Alex’ is much more rigid about how much time they use to complete the task before them:

*“... because of the relatively short turnaround time it proves more difficult to actually spend time marking but I’m a really sad person I actually time every assignment I mark ... 45 minutes for a 2,000 word assignment ...”.*

Whereas ‘Lesli’ discusses workload and the turnaround time for marking and moderation at their institution:

*“... The module yes ... I think I marked thirty scripts ...which isn’t unusual for this place. We have a fortnight to mark them in and we have the moderation **[fifteen working day turnaround – feedback to be returned to the student]** ...”.*

Another interviewee, 'Mel' confirms how they had to complete the marking of 30 assessment artefacts within the same timeline, without any cognizance of their learner status in relation to marking and grading. At the same time they have advice for other new lecturers:

*"Plan ahead in your diary ... make a point of finding out what you are going to be marking ... in advance ... because even before the marking comes in you need time with students to support them [to understand what is required from the assessment]"*.

The point here seems to be completion of a marking and grading task requires preparation, especially if the lecturer is new to the subject or the task itself. It appears 'Mel' was concerned about the preparation new lecturers receive for their marking and grading role. This interviewee considered that such nuances may not be taken account of, in policies or guidelines available to give advice and information about assessment.

The way 'Lesli' approached the workload seemed to stem from a previous experience:

*"... I learnt to protect myself. I was burnt out once and you never go back there again. So I learnt then to protect myself and make sure I allocate the time. We have our assessment schedules so I protect time in my diary for my marking ... I protect time for the initial scripts and I put time in for the extensions and referrals. Now if we don't get any then I have got space in my diary"*.

The sheer size of the workload and the pressure of when it was required also impacted upon the sense of fatigue 'Billie' felt:

*"If you are really tired and you are thinking I have got to get these done by tomorrow because I am so far behind. That kind of notion"*.

'Billie' is able to adeptly express why a situation like this frequently occurs:

*"...it is difficult when you are trying to slip marking in between a number of bits and pieces I think"*.

'Lesli' believes that the best way to approach the workload of marking is to view it as:



*“a job to do and I see it as that ... it is sometimes a bit of a slog. I see it I owe it to my students the marks that I give ... But then you get the pressure of teaching. And the pressure of your links [**clinical practice areas - link lecturing**] and the pressure of everything else. Your module leadership all these other deadlines coming. So you see this marking time eroding so you do it in your own time”.*

‘Gerri’ comments on something similar where it has become an accepted practice for using non work days to complete marking:

*“... I read them maybe five at a time have a break so it will maybe take me if I have got a few days to do it I won't do them all in one day but I'll do four or five one day four or five the next and probably over a weekend a Friday to a Monday ...”.*

Two interviewees discuss the impact of a deadline and the effect on them individually. ‘Eddi’ considers deadlines are relative to what needs to be completed:

*“... that depends, if you have not got a large amount to do its fine, if you have its a huge challenge ...”.*

Whilst ‘Harri’ reviews the concept of deadlines as impacting upon their marking and grading practice, especially when considering assessment artefacts which may not be reaching an appropriate standard:

*“... I usually go back on [**that artefact**], not all the others, haven't got time for the rest of them, its time that's the problem ...”.*

It is interesting that only one interviewee, ‘Danni’, portrayed their frustrations with the current workload and time limits

*“You know we can't do it any quicker. Sorry but we can't not if you want a rigorous process”.*

This opinion is possibly shared by their colleagues, in relation to workload and the drive to prepare feedback on written assessment artefacts at an increased pace. This interviewee appears to wonder if the task of marking and grading of an assessment artefact can be accomplished thoroughly with objectivity and accuracy in the current work regime.

The question of time limits becomes a stark reality when considering the position of 'Lesli', when it is not unusual for them to be faced with 30 assessment artefacts, each 3000 words long, when they have only set aside three or potentially four days to mark all of those papers. Interestingly another interviewee 'Danni' highlights they can only mark and grade three or four assessment artefacts in one day. At that rate this batch of 30 scripts would take between eight and ten days to complete, more than double the time 'Lesli' has allocated to complete the task.

Finally 'Harri' brings all of this into perspective by considering the whole process, and the implications of the marking and grading process:

*"... I hate it cos I think there's so much rides on it and I just don't feel confident so I do take probably more time than I should do on each one because I just don't feel, I never feel confident about the **[mark or grade]** ... It's just too important, it's too important to the students".*

This sub theme has explicitly confronted marking and grading time limits and work load. Many institutions at the request of students have instigated shorter and shorter deadlines for the return of feedback, which from evidence provided by interviewees has an impact on their role as a whole. This sub theme has facilitated the exploration of this component of contemporary lecturer practice.

#### 4.3.3.g MESSINESS OF MARKING AND GRADING

For five interviewees, undertaking and then completing marking and grading of assessment artefacts is an untidy affair hosting many areas for disagreement or raising uncertainty for those involved. An example of this is demonstrated by two of the interviewees 'Alex' and 'Gerri'. Both identified the difficulty of allocating a precise mark to an assessment artefact, when the scheme used at their institution called for use of a full range from 0 – 100%. They individually discussed the difficulty in being able to differentiate clearly between artefacts which were say to be awarded 52% rather than 55% for example. 'Alex' considered that they had taken the initiative and presented marks for a group of students which were based on allocating marks in 5%

increments e.g. 50%, 55% or 60%. On recounting their experience at a board of examiners some years before, 'Alex' describes the Chair's reaction:

*"... he saw this and said this is a bit odd ... and although he agreed with me he said that it must never happen again ...".*

Whilst 'Gerri' who worked at a different institution also talked about the process of allocating a mark after having evaluated an assessment artefact:

*"... I really don't think the difference between 55% and 56% makes any difference to me If I want to indicate something is clearly an A grade I will give it a 75 or 80%, if I want to clearly indicate it is in the B grade I will give it 65% you know ...".*

At that institution the use of the percentage scale can appear to lead to problems. The use of grades can become as problematic, when there is the introduction of -ve or +ve as in for instance A- or a B+, it could be argued that those using grades would have the same level of difficulty highlighted by both of the interviewees above.

'Gerri' also disclosed their approach to use of the available percentage scale:

*"... I tend not to give 61% or 59% or small increments like that because I just find that that leads one to arguments, you know the students appeals people on the **[exam]** board you know, **[by saying]** you weren't clear, did you mean this or, or whatever so I try and make the job of the exam board and the student interpreting the mark as easy as possible, so I tend to just do it in those gradations **[5% increments]** ...".*

The difference here being there were no complaints or mooted dissatisfaction from more experienced colleagues or examination board Chairs. It probably helps that this interviewee has had 25 years' experience in health professional education. In the cases of 'Gerri' and 'Alex' their approach was one in which they wished to remain in control of the mark or grade awarded and wanted to transmit a clear message to all concerned as to their thoughts in relation to the value of the assessment artefact.

'Gerri' further explained their rationale for this approach, even whilst acknowledging that:

*"... I know people can be critical of that [the above approach], but when you are not dealing with something that is like chemistry or science or mathematics where you can assign a specific mark at the top, then I think we are kidding ourselves to think we can give a very fine grain mark ..."*

This fits with the approach 'Alex' was trying to achieve, but was then criticised for, they said:

*"... I don't think you can mark terribly accurately, I thought putting them into range or putting them in a bracket was as accurate as you can be ..."*

This interviewee is voicing the opinion that there is an inability to mark or grade with an accuracy, exactness or precision that the awarding of a mark or grade gives an impression of. They are instead calling for something more akin to the notion of validity which has about it an idea of soundness, reasonableness or justifiability and that this could be more effectively achieved by placing the mark or grade within a bracket. Nonetheless this does not account for what those marking and grading would do when they considered the artefact they were evaluating is at the edges of the proposed brackets or boundaries.

Difficulties appear to arise for individuals for unexpected reasons; 'Mel' discusses their approach to constructing feedback on the evaluation of an assessment artefact and the reaction of their colleagues:

*"... some people say don't write a lot of feedback because it reflects badly on the other lecturers who don't write a lot of feedback ..."*

'Mel' as a fairly new department member was steered to conform to the lowest common denominator, being a provider of limited written feedback comments on assessment artefacts to ensure their colleagues approach was not criticised.

This then leads to the next area of messiness in relation to marking and grading and that is how criteria are used to evaluate the assessment artefact. Use of the criteria concomitant with lecturers' desire to award a mark or grade which fits with their interpretation of the assessment artefact, when on occasion, evaluation with the rubric seems inaccurate. 'Danny' illustrates this:

*"... I am not saying I don't ever sometimes I get a mark and I think oh that seems really high and then I go relook at my comments and where and I often find I have ticked the wrong box or something but I do try to criterion mark ...".*

'Mel' describes their process very clearly:

*"... and then I look and I don't just check one against another I sort of check one against all the others so it is usually that one that I change and I think is this one better than all the others then it deserves. Or I'll get one and I will think this is the best one of the entire class and I have given that one seventy five well unless I find one that is better than that, that one is my **[benchmark]***

In using this process it is not unusual for 'Mel' to end their evaluation of the assessment artefacts with achievement, for the student cohort, forming a bell curve, the curve of normality which is seen in standard data. When asked whether this was something which would be found in their student achievement data 'Mel' said:

*"Yes. Definitely in mine you would. Definitely".*

'Lesli' on the other hand highlights a further issue, related to the elements of the marking and grading criteria:

*"Sometimes you are marking this work and you are looking in the marking criteria and you can't find it **[the phrase or term you want to use to describe or categorise the work]**. So sometimes that does happen then what I do is I put a comment on the bottom about it".*

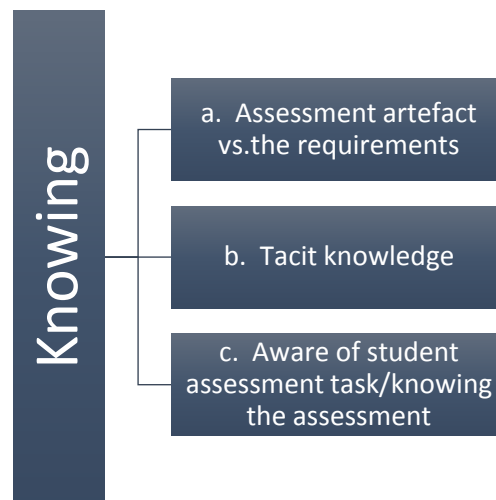
This occurs in spite of that institution having marking and grading criteria which are specifically developed for each individual module assessment. However, when lecturers need to employ a number of strategies for using the

marking and grading criteria it is not difficult to understand the complexity encountered. For instance 'Mel' highlights the issue of having had no involvement in a module but then being required to mark or grade assessment artefacts. They discuss some of the pros and cons associated with this task:

*“... I think one of the modules did produce a little bit of guidance ... like a breakdown of what they would expect to see in that section and examples of what is good and what's not. But a lot of them don't provide anything whatsoever and you might not have taught on that module you might not even be involved on that blackboard site so you don't even know what they have been taught”.*

The description of this interviewee presents a picture of a lecturer being required to complete the marking and grading task for an assessment artefact of which they have no knowledge or involvement in the teaching input. 'Mel' appears to hint that it may be possible to complete the task with some level of confidence when guidance is provided. It appears that this is not an expected component. This interviewee highlights an interesting possibility and that would be the difference in evaluation between those who know the module and the lecturers who do not. In essence it may not be possible for a valid evaluation for the assessment artefact to be achieved.

#### 4.3.4 KNOWING



**Figure 9 CI Major Theme**

This major theme brings together aspects of a concept I have called knowing. Here this is related to knowledge of the assessment and goes beyond awareness raising but includes how such knowing influences the process of evaluating the assessment artefacts.

##### 4.3.4.a ASSESSMENT ARTEFACT VS. THE REQUIREMENTS

This sub theme was explored by interviewees expressing what they saw in assessment artefacts which were both successful and not successful at the assessment stage.

‘Alex’ mused at how an assessment artefact could fail to miss the requirements on so many levels:

*“... there was very little thinking behind what they were doing ... they had quite obviously not read and taken in the quite good information about how to construct their assignment ... they weren't doing things ... they weren't using sub headings they weren't using the Harvard system correctly and I think ... the big thing they were doing wrong was that they hadn't read round the subject”.*

‘Mel’ concurs with ‘Alex’ and their position in relation to how assessment artefacts are constructed:

*“... yes it were quite clear the amount of time and effort had gone into it and the amount of reading ... short reference lists that were all NHS Choices and Talk to Frank and Stoptober and stuff. Really under the word count. Very superficial and you could tell who had made an effort ...”.*

Whilst ‘Chris’ discusses how efforts to support an improvement in written assessment artefact had not been acted upon:

*“... there were things that I had mentioned, things that we had talked about, advice given that hadn't necessarily been taken on-board ...”.*

It would seem the notion of a written assessment artefact that does not meet the requirements of assessment continues on into the resubmission or supplementary period (after an unsuccessful first submission):

*“... they don't understand that it is just not one thing that makes the one percent it's the whole thing...”.*

This was an important position to consider for ‘Chris’ as they allocated marks or grades for an assessment artefact. They felt it inappropriate to give the wrong sense of what corrections would be required to increase the assessment artefact sufficiently to one which was adequate enough to achieve a pass mark at re-assessment. This fits with their stance of never awarding a fail mark which is only 1% or 2% below the required pass mark of 40. ‘Chris’ feels that this gives the idea that only minimal changes or improvements are needed to secure success. ‘Billie’ concurs with this perspective of assessment artefacts which are found wanting, the interviewee on reviewing such artefacts considers:

*“... you go actually you have not even looked, you have not questioned this [**you've simply described**], you have not done any comparing and contrasting [**an essential element of analysis**] ...”.*

Even ‘Harri’ explores how they may review an assessment artefact carefully, if all the criteria for a pass is not evident at first glance:

*“... I think what I'm trying to do there ...is to, is sort of see is there sufficient information and if there's sufficient information but it's not, not written particularly well then I'd be looking at a borderline pass ...”.*



'Billie' also identifies how they may approach an assessment artefact, which on the surface may appear to be lacking:

*"... well it is kind of isn't it. So it is a layer of it, actually you have to dig a bit deeper..."*

Whilst it initially appears 'Billie' considers there is a lot of information, by saying 'dig a bit deeper', they are referring to being able to decipher writing that is less articulate and hence more difficult to understand and select the evidence of successful achievement.

#### 4.3.4.b TACIT KNOWLEDGE

'Gerri' highlights an issue which is implicated in making a judgement by those evaluating assessment artefacts and that is the tacit knowledge of the assessment itself that is held by the lecturer:

*"... If you set the question you have no excuse not to know what you are expecting, cos you should not only have criteria in a general sense but you should have criteria in your head of what you want out of this ..."*

This indicates that 'Gerri' considers knowledge of what an assessment is about, is located within the lecturer who designed and prepared the assessment task. They also then go on to consider what would occur if the individual in charge of the assessment task was not the individual to complete the marking and grading process:

*"... if you are marking on behalf of somebody else then it is a bit difficult a bit different isn't it ..."*

'Gerri' here begins to explore the impact of this lack of implicit or personal information of the assessment artefact.

'Lesli' talks of an assessment artefact and how they would interrogate it:

*"... partly memory and partly understanding how it is written and how it can be written ..."*

Here it is evident that having some prior understanding of what is expected from the artefact supports its evaluation by those marking and grading.

'Billie' goes a little further than this by insisting there is a particular type of:

*"... propositional knowledge ... it is you know how, but you can't always explain how, like riding a bike. You know how to ride a bike but you can't always explain to someone how to get the balance right. Cos they have to do it themselves ...".*

For 'Lesli' this knowing how comes:

*"... from past experience and this is where it comes in you have got past experience when you have marked really excellent pieces of work and you know what an excellent piece of work looks like ...".*

It seems that is one of the most important elements being suggested by interviewees gaining experience by contact with assessment artefacts. Along with this insight into the assessment comes through involvement in the design and compilation of the task which serves to increase the lecturer's knowledge of the requirements. The level and type of knowledge discussed by interviewees in this sub theme is developed through implicit rather than explicit means, personal contact and transfer of tacit knowledge; learning occurs though not via formal instruction.

#### 4.3.4.c AWARE OF STUDENT ASSESSMENT TASK/KNOWING THE ASSESSMENT

As has been hinted at in the preceding sub theme, interviewees considered that it is imperative that a lecturer is aware of the assessment task and therefore what the completed artefact will look like. Interviewees talked about finding ways to make themselves familiar with information about the assessment artefact. 'Eddi' explores their own process, especially when they are unfamiliar with the module:

*"... the first thing I do is I will make sure I have read the criteria that's been given to the students so I know exactly what they are being marked against ...".*

'Billie' provides a similar reason for the need to review the requirements:

*“... I always don't trust my memory enough to remember exactly ... I print them off and I have them on the side ready ... So I read through them so I know exactly what I am looking for”.*

No matter how much marking and grading the interviewee 'Billie' undertakes they feel that it is almost impossible to become a "supreme expert". A number of extraneous factors may mitigate against this, these include having to re-read to check for understanding, reviewing the learning outcomes as it seems almost impossible to keep what is required in sharp focus. This perhaps reveals the fallibility of human beings, that we are not machines which can recall exactly what is needed in the exact situation, with consistency.

Both of these interviewees believe that making themselves familiar with the learning outcomes will guide them in the marking and grading of the assessment artefact. Whilst 'Harri' knows they understand what is required for the module and is familiar with the learning outcomes:

*“... I didn't have to go to them as I know what's expected. I've done it that many years ...”.*

'Harri' keeps their own written notes of observations whilst reading the assessment artefact, then in conjunction with module learning outcomes the marking and grading criteria are used by this interviewee, but all with varying levels of importance. Interestingly they did not acquire the learning outcomes to check as they had marked this particular subject on a large number of occasions and considered that they had the information about the outcomes engrained in their mind. This is a different perspective to the one held by 'Billie', who wanted to refresh their memory.

However, 'Mel' highlights another perspective that is worth considering they say:

*“... you have got a sort of a personal, sort of like a preconceived idea of what this assignment is going to look like ...”.*

'Harri' sends out an important message in relation to the ease or difficulties that need to be overcome by those marking and grading assessment artefacts. According to 'Harri':

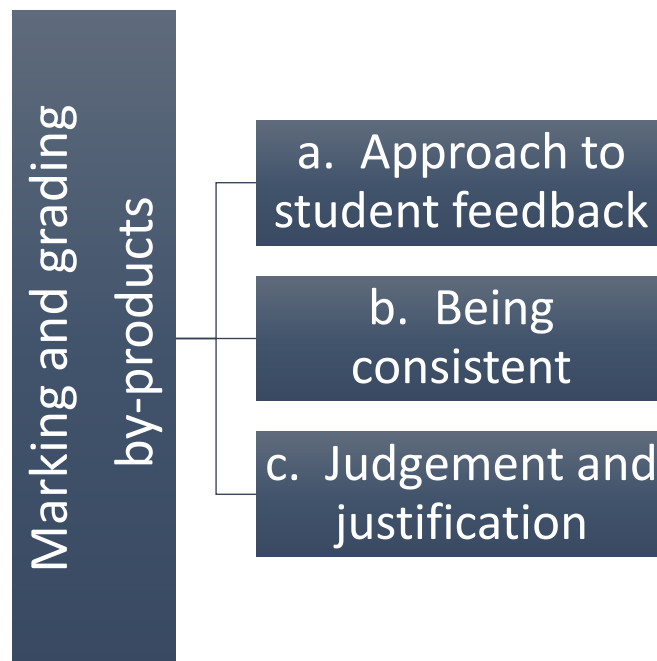
*"... If you're doing something that's fairly general or if you doing it from another field of nursing then it's about whether you can understand what they're saying. I do wonder sometimes if we're not just a little bit more picky because you know the subject and you know what you're expecting ..."*

Although directed by learning outcomes or assessment guides, a lecturer's knowledge on the subject area, the level of subject expertise and assessment artefact match to preconceived expectations could have a significant impact upon the final evaluation.

It seems lecturers marking and grading the assessment artefact are relying upon information produced for students, but in all cases this may not be sufficient to produce what is seen as a fair and just evaluation.

Preceding excerpts illustrate that lecturers depend upon information produced primarily for those undertaking the assessment to produce what is seen as a reasonable and objective evaluation, but this remains an imperfect system.

#### 4.3.5 MARKING AND GRADING BY-PRODUCTS



**Figure 10 CI Major Theme**

This final major theme deals with aspects of the marking and grading process explored by interviewees, which did not easily fit into any of the preceding sub themes. In some ways these sub themes could be viewed as an outcome from completing an evaluation of a written assessment artefact.

##### 4.3.5.a APPROACH TO STUDENT FEEDBACK

‘Alex’ mainly explores the potential of feedback to be focused on supporting the student to develop or improve their academic skills. To do this they express the need to ensure that any feedback is clear and understandable:

*“...I felt it was important for the student to know and be under no illusion what their mark meant ...”*

For this interviewee clarity seems to be the focus as they are keen to ensure that the mark or grade received by the student requires no translation or interpretation. They craft the feedback so that it is well-defined and provides an effective point from which to focus student development i.e. they will know what more they need to do:

*“... they need to **[to be able to]** move on from that, because what is forty two percent this year could be thirty two percent next year...”*

For this interviewee feedback on an assessment artefact should act as a springboard from which to develop future work, if the feedback is constructed in an accessible way for the student.

'Billie' is keen to communicate their approach:

*"... for me feedback is really really important and I give very detailed feedback and so I am thinking about the ones I couldn't tell you what the grade was but I would be able to recall where the feedback elements were..."*

It seems for this interviewee the crafting of appropriate feedback is the most important aspect of the marking and grading process, holding more significance than the mark or grade awarded. 'Billie' appears to place a great sense of importance on this aspect spending time carefully considering the qualitative comments. Interestingly the interviewee did not make evident how they expected the feedback to be utilised.

Another approach to providing feedback that this particular interviewee explores is the use of annotations. 'Billie' states:

*"... very happy annotating scripts as feed forward information to people and then it is really helpful particularly to people who fail to get it, look this is where the issue is on this particular piece ... this is what we need to look at and this is how I need to enable you..."*

This extract presents evidence of this interviewee anticipating that feedback written on the assessment artefact could be effectively utilised to make improvements, by the student, going forward.

The style 'Chris' takes in their approach to constructing feedback for the student is as a result of making notes:

*"... so I do scribble because then when I come to do the feedback sheet I can then look and see what ... I need to make comment on. Cos memory is an issue with me..."*

These 'scribbles' are used by this interviewee to assist them in constructing the feedback for the assessment artefact. The approach of writing notes to improve the final decision making is also explored by 'Gerri'.

'Billie' goes on to reveal the driver which appears to underpin their approach to student feedback:

*"Well we are supposed to be enablers ... and part of that is enabling learning how to do this, cos I didn't know".*

Their own experience of feedback and having a limited understanding of how to construct a critical argument culminates in the need to provide examples for the student.

A further extract from the interviewee 'Chris' highlights the ways in which they consider they try to encourage students to develop their writing, by reading and by closely exploring texts they read. They use the analogy of:

*"... see[ing] it like a picture they see a piece of art on the wall and you say oh well that is a picture of a house. But what you don't know is well is it oil or is it watercolour. Have they done it with a brush have they done it with a spatula, is it on canvas, is it a print. You know that kind of thing. All the bits that have put it together that make up the difference between one picture and another ...".*

At the institution where interviewee 'Lesli' completes marking and grading an interesting situation has come into existence, they are in the middle of implementing electronic artefact evaluation and as such are running two systems. A number of lecturers continue to print copies of the assessment artefact which they annotate, but an electronic feedback template is issued to the student:

*"So I annotate the paper scripts and keep the paper scripts for a little while afterwards so that if they want to come back to me I can pull these out. They don't get the paper scripts back. No it is all electronic ... **[if they are electronic they get access to the annotated artefacts online?] Yes yes".***

These interviewees see the provision of feedback as a way of opening a type of dialogue with the student. The view here is the provision of feedback as a way of encouraging the adoption of good academic habits that will lead to improvement of the quality of the assessment artefact in future submissions.

#### 4.3.5.b BEING CONSISTENT

In the context of the interviewee extracts presented here, being consistent can be related to applying the rubric for marking and grading. On the other hand for 'Alex' being consistent was related to their perception of the value of the assessment artefact not reliably matching with the rubric mark or grade:

*"...I will look at the mark and the rubric ... and think this is not consistent with what is in my head and then I have to have a little debate with myself about which is right ..."*

'Billie' has a similar take on the matter, with their approach to consistency:

*"...like, [I] have a debate about actually I think this should be higher or should be lower ..."*

This would occur when 'Billie' considered they were trying to be fair and consistent to an individual assessment artefact, as well as within a group of scripts.

For 'Mel' the inconsistency occurred through unclear or ambiguous information provided for preparation of an assessment artefact:

*"So it didn't actually say pick a patient pick a model but it clearly insinuated that they should ..."*

It could be argued here that this interviewee knew what was missing from the assessment information, and this shows they had an in-depth understanding of the assessment requirements, over and above what was formally available.

'Alex' makes the point that consistency is difficult to achieve when all the individuals involved in the process of marking and grading have their own ideas about what they are expecting from an assessment artefact. They consider that the use of an evaluative tool like a rubric:

*"... tries to bring objectivity to the process and at least you have got one standard set of criteria there that everybody is working with but that doesn't stop me as an individual being different to another person's individual but at least you have got some sort of common ground which is clearly explicit..."*



Being consistent has a number of different applications by interviewees in these findings, but essentially this is a way of speaking about the concept of fairness.

#### 4.3.5.c JUDGEMENT AND JUSTIFICATION

When interviewees discussed how they approached coming to a judgement about an assessment artefact they broached it in a variety of ways.

For instance 'Danni' talked of having a "gut feeling" they explained this by saying:

*"... I think you can tell by the way it is written, if the language is academic if it is coherent if it kind of makes sense ..."*

'Alex' said that their final evaluation of the artefact does not rely solely on initial feelings that they experience, but that they will also use the tools available to them:

*"... I need to go on to say in my defence ... that I then go to marking criteria and fill that out and see what that comes up with. Cos that is more objective to me for a start but subjectively ... when I am reading through what I am looking for is are the is the student understanding the question is the student does the student demonstrate enough evidence that they are understanding the subject matter and then is the student relating the subject matter to **[to their speciality]** ..."*

'Billie' goes on to highlight the difficulties faced by the lecturer when having to make a final evaluation on the quality of work before them:

*"... Sometimes people are very descriptive but then they come up with an absolute gem ... You think it does not quite fit. So then you have to make a judgement call and be as fair as you can be ..."*

A judgement is made by interviewees on awarding marks which are two percent rather than one percent below the designated pass mark as being more effective at demonstrating poor achievement.

'Chris' agrees with this approach, saying:

*“... I don't think it does the students any favours giving them a thirty eight or a thirty nine cos all they read into that is I have got one ... [or] two percent to make up and they do something like reorganise the paragraphs or write a bit of a longer conclusion ... they think that is sufficient to make up the difference. And it is not and they don't understand that it is just not one thing that makes the one percent it's the whole thing so I avoid those sort of just under the forty marks because I don't think it helps the student ...”.*

Interviewees actively avoid awarding a mark or grade which they consider would not offer enough information to facilitate amendments by students to assessment artefacts which are of the correct depth or breadth. This judgement does involve a certain amount of review on the part of the lecturer, to enable them to express more certainty in the mark or grade awarded. While this may initially be based on a 'feeling', 'Chris' for instance sees that:

*“...like I say, a feel that this is just a pass or this is a fail. So depending on that I'll go back and have a look at the marks and often go back and look at sections of the assignment that I think well can I find something ... you know go back and check myself out ...”.*

It seems from reviewing these excerpts that judgement is not made as a finite thing, but something that can be seen as fluid until the lecturer either confirms or refutes their initial findings. As well as confirming when an assessment artefact had not reached the required standard, there was also a feeling (expressed by 'Chris') that they had an obligation to express the things done well to attempt to maintain the self-esteem of the student.

Three interviewees made specific reference to being able to justify the marks or grades awarded for an assessment artefact. For 'Terri' they more or less felt that their use of the marking and grading rubric clearly supported the mark or grade awarded:

*“... maybe use that as a tool to inform your mark and to more or less justify your mark if you get questioned on it from either staff or students if they are coming to you for feedback ...”.*

'Terri' then goes on to qualify their rationale for providing feedback, which was deemed by some of their colleagues to be excessive:

*"... I tend to give too much feedback ... so that the students improve and they get this kind of formative feedback like feed-forward to other assignments. But also its selfish from the point of view that I know that I can justify to that student why I gave the mark I gave ..."*

The level of importance in which this is held by 'Harri', is revealed in the following excerpt:

*"someone else might mark it and give a different mark I appreciate that but I think on the whole I can put hand on heart say that I can fully justify why I gave what I gave..."*

The aspect of justification appears to be the method by which those involved in marking and grading are able to articulate the decisions they made, increasing the confidence in those decisions. The process of justifying their judgements explores a different perspective of arriving at a final mark or grade, enabling lecturers to attach a rationale to their decisions.

#### 4.3.6 SUMMARY

This interview method facilitated the gathering of data in relation to what was remembered about approaches to and the execution of marking and grading by interviewees, who were experienced and novice lecturers practicing in HE. Five major themes were identified and are explored through concomitant sub themes providing a window into the world of the lecturer completing marking and grading. The method for reducing the data into manageable proportions was described at the end of Chapter 2 and again here at the beginning of this chapter by the application of framework analysis. Decisions made about which data to display to illustrate my findings became a reflection of me as a researcher, this is explored further in Chapter 8, Reflexivity (8.2268) and is related to my ontology and epistemology espoused in Chapter 2. Interaction with the data as both a lecturer and researcher facilitated in-depth understanding and interpretation, each shaped by my unique experience. How this data relates to the

contemporary knowledge on the subject of marking and grading will be explored in depth in Chapter 7.

The following Chapter (5) presents the method for the second interview technique used for data gathering in this thesis, Protocol Analysis.

# Chapter 5 PROTOCOL ANALYSIS INTERVIEW METHOD

*"People are like stained-glass windows. They sparkle and shine when the sun is out, but when the darkness sets in, their true beauty is revealed only if there is a light from within."*

Elizabeth Kübler-Ross

This chapter explores more closely the second of two data collection methods used in my research study. It begins with a presentation of the origins of this approach, including identification of some of the benefits whilst acknowledging potential limitations recognised in the literature. Close attention is given to the interview process leading to data collection, along with a consideration of the interview setting. Information relating to the lecturers participating in Protocol Analysis interviews is presented. The approach for data analysis is presented in Chapter 2. At the end of this current chapter, an approach to increasing credibility and dependability of findings is presented. Finally, the chapter culminates with a summary of this data collection method.

## 5.1 WHAT IS PROTOCOL ANALYSIS?

Protocol Analysis (PA) was conceived in the 1980s by Ericsson and Simon (1980) to gather information about the mechanisms and cognitive processes occurring in research interviewees during the performance of a behaviour or task. Collection of this information is done through asking the interviewee to verbalise their thoughts, which are recorded in a concurrent approach. The interviewee is expected to relay all new thoughts as soon as they are generated and surfaced in their attention (Ericsson and Simon 1993: pg xiii). They are asked not to describe or explain their thoughts but to simply relay information in a verbal format as soon as a new thought enters their attention. It is the concurrent nature of this data which elicits information that would normally remain unavailable to researchers as study outputs.

Using verbal reports data is not new, it has been consistently used by researchers as evidence and in many approaches interview data is most frequently retrieved in an asynchronous fashion. This is achieved by conducting a post-performance interview where interviewees are asked to relay their thoughts, focussing on the execution of the behaviour or task being examined. However these thoughts need to be retrieved from the interviewees long term memory and involve retrospective probing to recover information required (Ericsson and Simon 1993). However when the PA method is enlisted, data is gathered from the interviewee as soon as it is generated.

Protocol Analysis, it seems, has the ability to establish and communicate an individual's cognitive processes and in the case of my research, in relation to the activity of marking and grading. Kuusela and Paul (2000) assert that the conscious act of verbalising all thoughts may have an impact upon the final object of attention. This occurs because a level of interpretation would seem necessary to change the thought into a word that can be communicated and then reliably understood by another person using the same language. It is suggested that we do not exist external to a world in which social activities occur in the absence of a means to describe them (Whorf et al. 2012). The language and the activity are mediated by the culture of its origin, with each giving existence to the other.

## 5.2 PROCEDURE

In my study, Protocol Analysis (PA), or more accurately described as Talk Aloud, facilitated the capturing of interviewees' responses to student written assessment artefacts, being gathered in 'real time'. In this context it enabled the exploration of the marking and grading of a student essay, through revealing the cognitive thoughts and processes occurring in the mind of the lecturer completing this task. The potential for this technique to illuminate problem solving and decision making has been previously demonstrated (Benbunan-Fich 2001; Funkesson et al. 2007; Crisp 2008). Therefore PA was used to record all utterances of the interviewee, as they were implored

to 'Talk Aloud', to speak all occurring thoughts, whilst undertaking the task of marking and grading a student essay style artefact submitted for the purpose of assessment.

### 5.2.1 CONDUCTING PA INTERVIEWS

Interviews were conducted in an authentic environment, this consisted of the interviewee having access to written assessment artefacts that required marking and grading within a specific time period. The PA interview took place in a setting chosen, to some extent, by the interviewee to reproduce, as far as possible, genuine approaches to marking and grading. All interviews were recorded using a digital voice recorder. Some further observation of PA interview interviewees occurred, when as the researcher I sat in the same room during completion of the marking and grading. I was able to observe non-verbal responses to marking, including gestures and facial expressions, these were observed as part of the curiosity toward interviewee reaction to assessment artefacts.

Capturing lecturers whilst in action is an appropriate way to illuminate their experience of and response to marking and grading of written assessment artefacts, illustrating Schon's (1987) 'Knowing-in-action'. Schon's theory on professionals and reflection, is one which demonstrates the practical knowledge held by a professional. In the case of my research the knowledge of marking and grading practice is demonstrated via a process of talking aloud, turning cognitive thoughts and processes about the practice, into language.

A post PA debrief was undertaken with all 11 interviewees. This was included to provide an opportunity for interviewees to take stock of the process they had been engaged in. Normally marking and grading is undertaken as a solitary process and I wanted to provide a chance for interviewees to explore their thoughts and feelings. The post PA debrief took the shape of a semi structured interview, with a group of 5 questions being used to guide a short discussion (interview schedule PA, Appendix 7). Freedom was given to the interviewees to ask questions at the end of the

interview for clarification of any areas that the interviewee wished to explain or to have explained to them. Retrospective verbalisation and reports were additionally included by Ericsson and Simon (1993) to aid in increasing the completeness of verbalization of the cognitive processes. However, in my research the focus was mainly to provide support to interviewees.

### 5.2.2 ADVANTAGES

Protocol Analysis (PA) necessitates that interviewees be engaged in an authentic task during which they are required to 'Talk Aloud', revealing a step by step journey of their thoughts and actions during the task. This 'talk' is thought to be representative of the interviewee's concomitant cognitive response to the task. In my research the authentic task is performing summative marking and grading of a written student assessment artefact, with the intention of exploring the real world activity of assessment via marking and grading. Employing this interview technique provides insights into the moment or 'very nature of the phenomenon' (Van Manen 1990). This approach facilitates the illuminating of the phenomena in a way which is consistent with the philosophical underpinning of this study, that of hermeneutic phenomenology. The lived experiences of lecturers, the hermeneutic perspective, is described from the logical standpoint of within and during the experience (of marking and grading), providing an important vista. This approach contends that PA permits an insight into a process which would not normally be possible and it does so from the perspective of the interviewee. This synchronous verbal data collection method has been used in a number of contexts, including psychology and information technology, specifically website design, testing usability, as well as in health and education settings (Aitken and Mardegan 2000; Benbunan-Fich 2001; Boyd et al. 2009).

This is a method which requires interviewees to verbalise their problem solving and assumes that verbal behaviour can be used for analysis purposes (Aitken and Mardegan 2000). Information gathered during this process may be viewed as a-priori, that is, the proposition is known independent of experience, and has no empirical foundation. The



interviewee has a belief in being able to perceive the truth of the information they are relaying during the PA interview (Baehr 2006). Despite information being generated as part of the marking and grading experience, synchronous data collection facilitates lecturers reflecting within the experience, a skill identified as “reflection-in-action” (Schön 1987). ‘Reflection-in-action’ is when an individual is able to observe what they do, whilst in the moment; this is knowing-in-action. The next phase involves reflecting on this knowing-in-action, being able to describe it and lastly to reflect on that description, potentially generating new knowledge (Schön 1995). This approach fits well with the philosophical and methodological stance taken in this thesis, a desire to uncover the lived experiences of lecturers marking and grading authentic assessment submissions, written assessment artefacts.

PA is the tool which enables these reflection-in and on-actions to be captured. Thus the technique is not thought to be introducing another step, rather it is surfacing what already exists by requiring these thoughts and actions to be expressed in a verbal and thus recordable form. These actions are implicit, characterised by the skilled practitioner portraying natural and instinctive actions, with their knowledge seemingly to be within these actions. Polanyi (2009) describes this as tacit knowledge, the know-how being embodied within the action itself. The embedded nature of this know-how can make it difficult to describe to others accurately and truthfully. This point is supported by Hycner (1985) who criticises the gaining of retrospective perspectives from interview participants, as information retrieved is unlikely to be similar to that of an individual from whom a description of the task is gathered whilst they are in the moment. Time is felt to be the mediator, that brings with it difficulty in being accurate about the exact thoughts that were related to an event, which is now in the past.

PA has the potential to reduce errors in reporting by not relying on retrieving information stored in the long term memory of interviewees. Information is gathered and temporarily stored in the short term memory and it is from here that reporting during an action will be done, thereby using the most contemporary information related to the action of the interviewee, without

further processing. This means that the interviewee should be able to provide minuscule, seemingly unimportant detail, anything which comes to the fore in the mind during the completion of a task. It is thought that access to this type of data facilitates collection of data which closely reflects the truth. However, Paul Ricoeur, a 20<sup>th</sup> century philosopher (1913-2005), maintains that any memory recalled can only ever be partial, interviewees are unable to retrieve every detail, therefore representing a version which is always prone to be fallible (Pellauer, David Dauenhauer 2016). This perspective is validated by Lee et al. (1999) when investigating the ability of parents to recall what vaccinations their children had received. Recall was not improved by the proximity of survey questions to the event, which provides evidence that the event was potentially not encoded correctly and therefore could not be recollected with precision.

### 5.2.3 DISADVANTAGES

When asking lecturers to verbalise their thoughts or cognitive processes, there is a theoretical potential that this could influence or interrupt the natural progression (Ericsson and Simon 1993) of the marking and grading task. This is more easily highlighted by briefly exploring what thought and thinking are. Thoughts are what are uttered during the PA process through the medium of language. For this to occur, it could be considered that thoughts are turned into the language to enable them to be uttered. Therefore it is plausible that PA will slow down the process being observed, marking and grading, in the case of what is being reported on here.

Interview techniques which do not occur in the moment of action or activity produce information through responses which are a-posteriori (Baehr 2006). Thus constructed reports are more likely to consist of interviewees theorizing about what they were thinking rather than their actual thoughts (Schön 1995). This type of reporting can be fraught with inaccuracies as many interviewees have difficulty in relaying what information processing occurred, by retrospectively calling upon their long term memory to undertake this reporting (Ericsson and Simon 1980).

Verbalisation of thoughts and thinking occurring whilst completing a task relies upon the interviewee having an awareness of all their cognitive moves. As an individual becomes more experienced, and thus, perhaps, developing expertness, actions required to complete a task become more implicit, necessitating reduced cognitive attention or effort. As defined in Chapter 1 expertness is demonstrated through the amassing of knowledge and skill which is acquired through experience and practice and not simply via time served on the job (Hoffman 1998). Familiar tasks can thereby be accomplished using limited conscious effort, interviewees can complete the task of marking and grading by remaining in the moment of the task, no examination or exploration of the thoughts are required as they unfold. If this is the case, then it is likely that some interviewees may not be able to complete a PA interview in the way that is required, to gain knowledge of the cognitive processes which are the first immediate thoughts that the interviewee becomes aware of during the completion of a task. Rather these individuals resort to describing or explaining what they are doing instead of uttering all the words which spring to mind. For instance some interviews would say "I am now going to review the reference list", rather than simply arriving at the page with the references and saying something like "I see they have used Harvard referencing, good and the list is in alphabetical order". The first example illustrates a describing of an action, whilst the second example shows the outcome of the action, being able to distinguish the referencing convention used. It is acknowledged that not all thoughts will be relative to the task being undertaken as it is easy to become distracted by mundane daily activities that simply pop into your head e.g. "did I take the soup out of the freezer?" However, interviewees are instructed to include all thoughts as they occur, such extraneous expressions were excluded from the final analysis.

It is feasible that the cognitive processes may be interrupted or indeed impaired by trying to gain access to knowledge that is usually inaccessible. The ability to perform the task may also be compromised by the presence of recording equipment and indeed the researcher (Kuusela and Paul 2000). In this case utilising PA may be damaging to the authentic task of marking and

grading, thereby impacting upon the output constructed by the interviewee. To reduce any such impact on the final mark or grade awarded to the authentic assessment artefact, ensuring that these artefacts are subject to scrutiny via internal moderation processes was vital.

#### 5.2.4 TRAINING

To facilitate verbalisation, the marker/grader was provided with training on how to verbalise their thoughts prior to commencement of the PA interview, as advised by Ericsson and Simon (1993). Asking the interviewee to verbalise, I considered, required a small investment in time to support the interviewee being able to reveal or develop appropriate skills. Training consisted of introducing several short exercises to exemplify the process of Protocol Analysis. Each interviewee was asked to look at a pre-prepared card placed in front of them, presenting a simple question (Appendix 8 p336). The exercises used were very similar replicas to those used by Ericsson and Simon (1993). Interviewees were then asked to talk through their solution, not simply provide an answer to the question. The thinking behind this was to ensure that interviewees had an understanding of what was required of them during the PA interview. Marking and grading is usually a silent, solitary process with a requirement to utter every occurring thought significantly diverging from this. Therefore warm up sessions were imperative to assist interviewees in becoming familiar with the practice of speaking aloud, verbalising their thoughts.

#### 5.2.5 SELECTION OF PA INTERVIEWEES

A total of 11 lecturers satisfied the inclusion criteria which included both novices, with less than 2 years' experience in Higher Education and more experienced lecturers. A major factor in deciding which lecturers could complete a PA interview was the availability of marking and grading during the timeframe proposed for the interview. When interviews were arranged, this factor was discussed with potential interviewees, with the interview being organised to facilitate the lecturer completing their marking task to meet pre-determined department deadlines for release of grades and feedback to

students. The existence of these constraints meant that potential volunteers were precluded from participating because of not meeting these prerequisites, or being unable to schedule a suitable appointment during the proposed data collection period with cognisance of pre-existing diarised commitments.

The type of student artefacts required were written assessment essays or short answer questions, rather than multiple choice papers which would lead to a reduction in perplexity as well as ambiguity in decision making. This is because there is usually an answer guide for multiple choice papers, with a right answer for all questions. Essay assessments require lecturers to come to their own judgement as to the quality of the answer on the paper they are assessing.

A recent research undertaken by Bloxham et al. (2011), completed a similar study with an interview sample of 12 lecturers from a number of different humanities disciplines. The differences between this study and my own, reported here, is the focus of interviewees who were from a variety of health backgrounds. Although my sample demonstrated more homogeneity than in previous studies, their backgrounds were from a number of different professionalised disciplines found in healthcare. This group of individuals by necessity continues to engage with particular discourses relating to patient or client care (Clouder 2005) or safety (Armitage et al. 2011) for instance promoting the notion of safe practice.

#### 5.2.6 MEMBER CHECKING

Triangulation of one PA interview was achieved through member checking of extracts and the themes to which they had been applied. This approach was adopted to provide assurance of the dependability of the research findings reported in this thesis. A discussion took place between 'Ali' and myself whilst reviewing the interview transcript. 'Ali' was provided with a copy of the entire transcript and then a document detailing each extract and the individual subtheme application. NVIVO software enabled the retrieval of this detailed data and presentation of it to support discussion with the

interviewee. Some exploration of sub themes occurred to enable questions from the interviewee to be answered through discussion and identification of researcher thinking. For instance 'Ali' asked about second marking and where this was considered, I was able to highlight the sub theme of 'Moderation' which includes all aspects of second or double marking, second consideration or blind double marking (marks or grades and feedback unknown by a second individual completing the evaluation task). This discussion satisfied 'Ali' and myself that I had been able to interpret elements from their interview transcripts appropriately, thus providing reassurance in the trustworthiness of findings as reported in Chapter 6.

### 5.3 SUMMARY

Protocol Analysis (PA) supports the exposure of interviewees' cognitive processes during task completion. In the case of my research this is related to the uncovering of lecturer response to written assessment artefacts presented for summative assessment. Utilising a PA procedure, interviewees are encouraged to verbalise all thoughts and reactions to the artefact before them. An effective PA verbalisation is one which demonstrates the immediate thoughts of interviewees and the outcome of these thoughts rather than a description of the action they may be undertaking. For instance an interviewee may review an artefact and verbalise the following: "... *this seems rather long, have exceeded the word count?*" rather than the following ... "*I am now going to check the word count as the paper seems to be rather lengthy.*" The difference being, in the second example the interviewee is describing what action they are intending on taking.

This is a method of data collection which has been used in previous studies in education both school based and higher education and thus has a proven track record in being able to elicit appropriate qualitative data. In my research study this method was selected as it facilitates close contact with the lived experiences of lecturers marking and grading, fitting with my research philosophy.

PA facilitates the gathering of data, real time, reducing the effect of fading memory and post task reporting. It is considered that this will elicit data which would normally remain hidden should a traditional post event interview be undertaken. Close attention to post interview debrief to facilitate the interviewee in exploring information which was brought into their attention, which may have otherwise remained hidden, during what is normally an individual and silent activity.

# Chapter 6 PROTOCOL ANALYSIS FINDINGS

*“It takes two of us to create a truth, one to utter it and one to understand it.”*

Kahlil Gibran - Sand and Foam

This chapter sets out findings retrieved via Protocol Analysis interviews as one of the methods of data collection and will begin with a brief presentation of findings from this approach. The data gathered was examined by utilising a modified framework analysis technique. The stages of the process as applied to this data collection method have been previously explored in Chapter 2.

The organisation of this chapter will be as follows: firstly, an inclusion of the research questions, next, a closer inspection of the interview cohort characteristics. The next section will pursue the themes developed through close engagement with the interviewees as they live the experience of marking. As was outlined in Chapter 5, all contributors completed an authentic task, whilst using the Protocol Analysis method for gathering data. Finally a short summary will complete this chapter reviewing the progress of theme development.

## 6.1 RESEARCH AIM

To investigate the process of marking and grading from the perspective of the marker. To recognise good marking and grading practice, deriving a model of ‘expert’ practice, aiding the development and proficiency of novice markers.

This interview method proposes to uncover answers to two of the three thesis research questions (2 and 3), whilst attending to the research aim, and maintaining a connection with the research objectives.



### 6.1.1 RESEARCH OBJECTIVES:

1. Examine assessment by exploring marking and grading practices.
2. Identify what are 'good' marking practices to develop a concept of 'expertise' in marking and grading.
3. Explore novice lecturers' thoughts on marking and grading
4. Explore cognitive processes and extraneous influences on marking and grading practice.

### 6.1.2 RESEARCH QUESTIONS:

1. What do lecturers believe they do when they are undertaking marking and grading?
2. What thoughts or cognitive processes occur in lecturers during the act of marking and grading?
3. What exists to ensure equity or consistency in professional or inter-professional marking and grading, this includes inter-rater reliability.

A return to the beginning of the intended study objectives is imperative, hence here a review of the research aim and questions, which serve as a reminder and will frame the presentation of the remainder of this chapter.

## 6.2 INTERVIEWEE KEY 2

Table 4 PA Interviewees

<b>PSEUDONYM</b>	<b>EXPERIENCE in years</b>	<b>INSTITUTION</b>
Maz	Less than 2	Post 1992
Toni	5-10	Pre 1992
Jae	1 ½	Post 1992
Shirley	2 – 5	Pre 1992
Ali	2	Pre 1992
Sam	5 – 10	Pre 1992
Jo	5 – 10	Pre 1992
Jacki	15 – 20	Post 1992
Bobbi	2 – 5	Pre1992
Nic	10 – 15	Pre1992

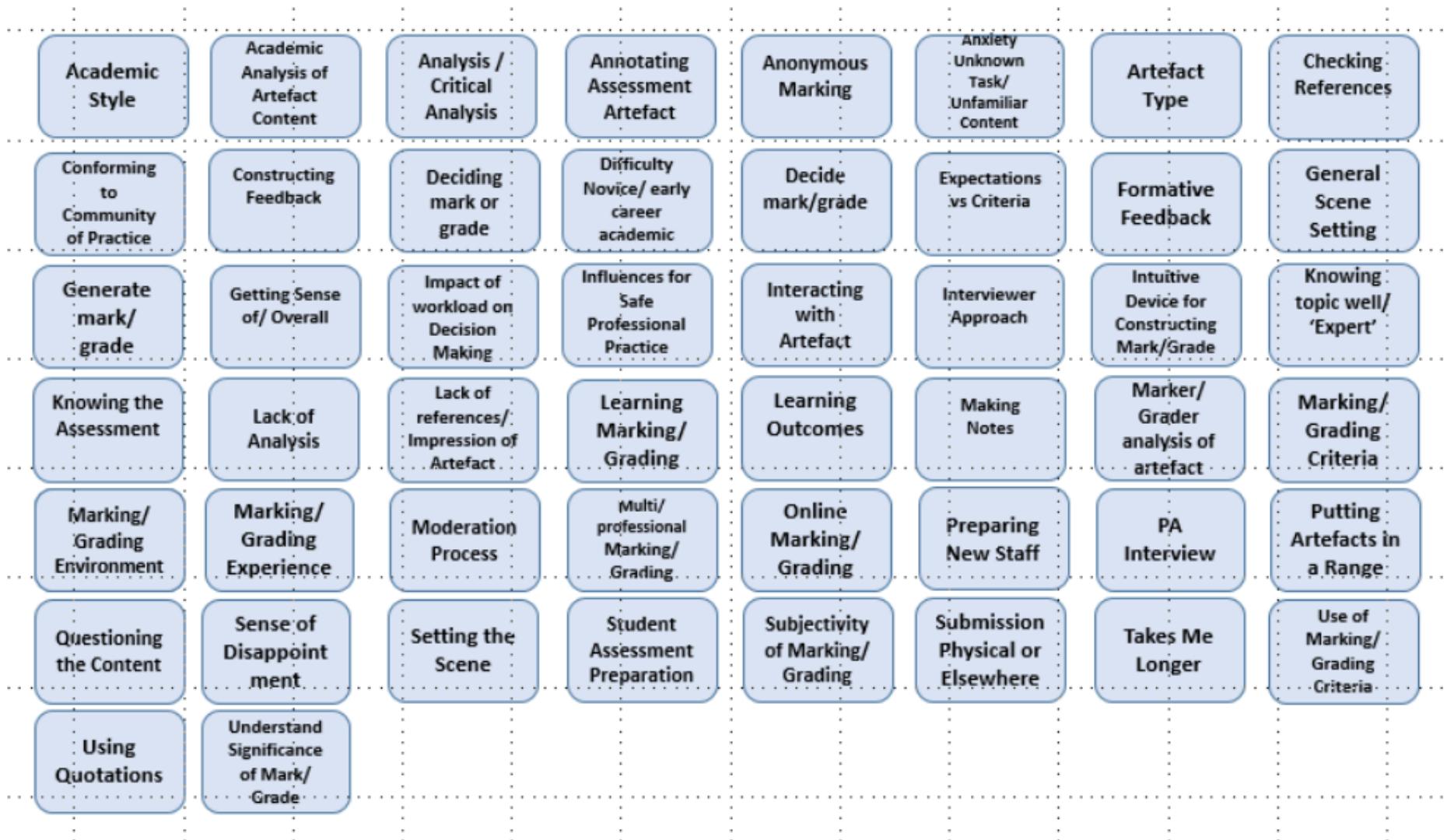
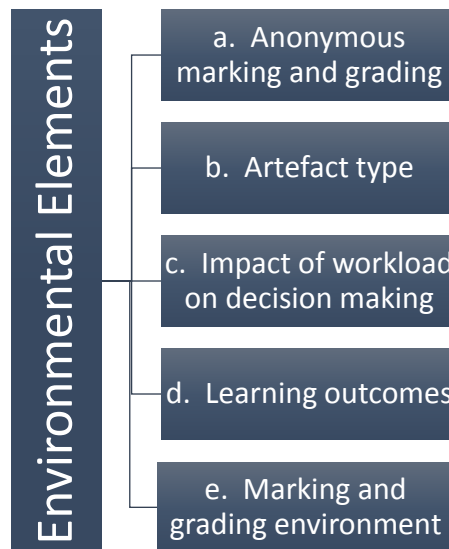


Figure 11 PA Initial Coding

## 6.3 EXPLORATION OF FINDINGS

In line with process highlighted in the methodology (Chapter 2), modified framework analysis has been employed to elicit a close examination of the data from individual interviewees. At the early stage in the process 50 codes were highlighted through reviewing the list of 'Nodes' (codes) in NVIVO (Figure 11). This list of codes could be viewed as a static list, prior to re-engaging with the data from which the codes were derived and then exploring links and similarities between them. These links resulted in a reduction of in the number of codes now seen as sub themes and then the creation of major themes. Each major theme was crafted from grouping together like sub themes, as I saw them, ending with a total of seven being created. The data to follow is presented by exploration of 6 of the major themes and the 27 embedded sub themes that appear to directly relate to responding to the research questions. The major themes are as follows: 6.3.1 Environmental elements, 6.3.2 Student assessment literacy, 6.3.3 Operational necessities, 6.3.4 Tangible marker or grader actions, 6.3.5 Implicit knowledge or actions, 6.3.6. 'Newby' issues. The seventh theme, Outliers, is related to the interview approach and will be explored via researcher reflexivity (Chapter 8.2).

### 6.3.1 ENVIRONMENTAL ELEMENTS



**Figure 12 PA Major Theme**

This major theme deals with issues that were raised by the interviewees which I have captured as factors which existed and were beyond the control of individual lecturers, these are viewed as extraneous influences. This major theme will be explored by way of five sub themes, which are related to the research objectives. Two other sub themes of online marking and submission of artefacts were identified but remain outside the scope of the research questions and objectives of this thesis and are therefore not explored here.

#### 6.3.1.a ANONYMOUS MARKING AND GRADING

Two interviewees expressed that they felt some difficulty in marking and grading an assessment artefact when they had familiarity with the individual who had produced it. 'Maz' sums this up well identifying that even when anonymous marking and grading was in place:

*"... because you actually supervise that group of students you can pretty well remember who has done what ...".*

The employing of anonymous marking and grading does little to prevent the identity of a student being revealed due to close supervision of the student workload. The lecturer is more easily able to recognise the content and

structure potentially revealing the identity of the student. This position is exemplified by the following excerpt, again from 'Maz':

*"...for me when I am reading this I can almost hear them saying it ...".*

The importance of this point is brought to the fore by 'Toni':

*"... because the person I've just read is XX I'm being judgemental, but I'm making an assumption that the next student in this list has a higher chance of written English not being her first language ...".*

This interviewee is in an institution which does not employ anonymous marking or grading and therefore has a list of students, they use this to ensure that they have completed all the marking and grading. This highlights the potential ethnic group of each individual to be assessed, as those names are revealed. The interviewee is explicit in identifying their own biases, in a sense they are going a step further which is to try to ensure parity by:

*"... make[ing] sure that you get an assignment in which a student is achieving similar mark and that English is their first language so that I can compare easily pieces of work ...".*

Another factor which impacts on the ability to undertake anonymous marking is the artefact itself, as 'Maz' illustrates:

*"... we do a lot of marking of practice portfolios within district nursing and that is not anonymous ...".*

This sub theme seems to identify potential benefits to maintaining anonymous marking; however, it is not without hard evidence that the ability to recognise individual student assessment artefacts is ever present.

#### 6.3.1.b ARTEFACT TYPE

Assorted assessment artefact types are explored here as a sub theme. The different approaches required to evaluate a variety of assessment artefacts which interviewees needed to navigate are presented.

Two of the interviewees referred to assessment artefacts as an assignment. For 'Jae' this took the guise of them trying to identify what the student was doing:

*“... So she has put here a little bit about what she wants in the assignment ... it's not set learning outcomes ... it is not in the module handbook and I am sure she has got it somewhere ...”.*

Whereas for 'Shirley' the assignment was much more discretely described:

*“...so she's mentioned the case scenario ... she's put that she is going to look at upper limb mobility ...”.*

Another interviewee 'Ali' was required to evaluate a piece of work described as a 'portfolio' but on closer inspection this incorporated a number of elements, increasing the complexity of completing marking and grading. The assessment artefact constituted:

*“... The portfolio I am marking today is ... at level six ... kind of consists of three summative pieces of work. One of them is a case study 2000 words where they have to reflect and justify a ... decision ... another one is a clinical management plan and another one is a formulary ...”.*

This interviewee would need to apply several different skills in order to evaluate the success or otherwise of these disparate elements submitted for assessment.

Two other interviewees identified that the assessment artefacts they were evaluating were based around the application of a critical appraisal technique. For 'Sam' this involved them in establishing:

*“...you could tell that the student had used the CASP to structure their work ...”*

This approach formed part of the guidelines which had been provided for structuring the assessment artefact. 'Jo' was reviewing the artefact they had come into contact with, with a different lens:

*“... she doesn't show me that they have got a clear understanding of the area, they have been asked to do a critical analysis of an article, but this is not what they seem to have done ...”.*

Finally, the assessment artefact which 'Toni' is appraising is challenging them to:

*“...keep thinking about are they doing what of been asked to do ... are they critically evaluating, have they understood what a critical evaluation is ...”.*

Further to this ‘Toni’ expresses another perspective and that is in relation to expectations:

*“... I was looking for them to critically appraise the ... protocol but what they've done is critically appraised the type of examination being requested, so I don't know, I'm not surprised if they didn't know exactly what they were meant to do because I don't think the module handbook is clear ...”.*

These last three excerpts illustrate that interviewees had to take time to examine the assessment artefact in detail to identify if it demonstrated achievement of the task set. Interviewees had to employ a number of different skills, knowledge and understanding not only of the artefact topic but also of the expected structure, related to the task that had been set for the assessment.

#### 6.3.1.c IMPACT OF WORKLOAD ON DECISION MAKING

Concerns of interviewees in relation to marking and grading workload and potential influence on decision making are illustrated. The following excerpts present different aspects of concern and impact upon marking and grading practice. The first of these interviewees is ‘Ali’ who during the interview expresses some unease:

*“... I am going to say this because I think it is important in marking that I think where a student is borderline, especially when the assignment is massive like this, I think a huge influence on markers is having to mark it again. And I know people don't say that but I really do think it is cos I just had it flash into my head just then. About like I've got to read this whole thing again ...”.*

This was a significant revelation as the assessment artefacts being evaluated by ‘Ali’ had been awarded a fail mark. Their initial concern was an initial dismay at the thought of having to interact with an, albeit, improved



assessment artefact for a second time. They then went on to qualify the approach:

*“... I am kind of aware of it enough to know that it wouldn't kind of effect my practice but I think it does for borderline. **[If you had 20 portfolios]** Yes absolutely. OK I do think its possible ...”.*

For another interviewee 'Jo' the issue of workload was explored through their approach to reviewing aspects of the assessment artefact:

*“... one thing that I have done in the past is to look through some of the references to make sure that they exist and that they are accurate in terms of the titles and things. The ones on here I know a few of them and they're fine I'm not going to check more widely. I don't have the time to check in that detail anymore, we've been pushed away from that. So it just comes down to making some comments, giving a mark ...”.*

A change in practice is viewed by this interviewee as having being enforced due to time constraints. 'Jo' explores this further:

*“... I think it's an extra challenge ... it's a time element and pressure to work you quite aware of time, trying to get through it quickly ... feel like I'm dawdling ... I am a slow marker ...”.*

The consequences of working at a pace which does not ordinarily fit with their approach 'Jo' states:

*“... I find it quite challenging to mark and actually it will tire me out, if I've done one or two papers ... they'll all blur into one yeah, its certainly challenging in terms of the time element ...”.*

The perspective of time constraints is explored by 'Maz' who presents the issue of quantity, as in quantity of assessment artefacts:

*“... also for me how many I do at once because I want to give number twelve on that list as much attention as number one ...”.*

The issue of quantity is eloquently explored by 'Jo' who simply states:

*“... and then I'll be quite confused as to what I've just read and whether I'm commenting on a paper two before ...”.*

It is then proposed that this confusion is counteracted by the individual marker and grader expressing their approach:

*“... sometimes feel that I need to go back to the first one because maybe feel as though I've not been fair and or actually you see **[different]** things as you go along ... and maybe ... that I've been quite tight ... maybe it seems to be quite good compared to the next so ones coming through ...”.*

The time aspect took on different connotations when ‘Toni’ considered how much time has been spent marking and grading an assessment artefact:

*“... I'm just thinking about how long it's taken me to mark this but I'm not worried because it's the first one I've marked and am reminding myself that I normally spend quite a long time thinking about what the assignments asked them to do ...”.*

This last extract highlights that time is required for in-depth reflection, facilitating exploration of the assessment artefact and what it represents. This action increases the time required to complete marking and grading tasks, when it can already be seen that achieving the right quantity of time presents a challenge.

#### 6.3.1.d LEARNING OUTCOMES

Each of the following five interviewees discuss referring to the learning outcomes at some point during the marking and grading process. For many of the interviewees this review took place prior to appraising the assessment artefact. For ‘Jo’ a review of the learning outcomes was a formal process because:

*“...they never change but I still like to have a look at them ...”.*

They saw this as an opportunity to refresh their minds on what their expectations would be for the assessment artefact. This appears to be an important or critical action to be undertaken and as ‘Toni’ highlights:

*“...so now I'm slightly annoyed with the module leader and starting to think again about how the information they have got to work from, but you know haven't got the module descriptor at hand or the module handbook or the assignment title ...”.*

The unavailability of learning outcomes makes completion of the marking and grading of the assessment artefact exceptionally difficult, as they have limited knowledge of the necessary content and expected structure of the artefact.

This theme is continued with 'Jae' who had to resort to reviewing their department's the virtual learning environment:

*"... I printed it off cos I thought they were on here, there is no learning outcomes in the module handbook which I find quite surprising ..."*

Where learning outcomes are available, interviewees made use of these in reviewing the assessment artefact and in establishing how to evaluate it.

'Jacki' explores how this occurs:

*"... So there are four learning outcomes and they are banded. I would allocate a mark to each of those ..."*

However 'Jacki' also points to the downside of simply using the learning outcomes to establish success or failure in the assessment task:

*"... and literally as soon as they **[student]** ... don't take that learning outcome up ... well it feels like a bit of a tick box exercise ... they miss one learning outcome not saying it happens but someone may fail them ... you just think that's not fair and that's not right. It is surely quality of information ..."*

'Jacki' then goes on to explore how the learning outcomes may well be applied:

*"... I have seen some of my colleagues will have the learning outcomes in front of them and our learning outcomes are too big for a lot of our modules ... ridiculous where you've got fifteen learning outcomes to one module assignment ..."*

Nevertheless for 'Jacki' there is a caveat that there are potentially too many criteria with which to evaluate the assessment artefact.

When the time comes to review the marking and grading criteria this may be when interviewees consider if they have the learning outcomes available to them, as is evidenced in this extract from the interview with 'Ali':

*“... Good coverage of content or learning outcomes met well. So the learning outcomes for the module are, I was going to bring a list, but I think I remember them ...”.*

However, as the evaluation of the assessment artefact continues ‘Ali’ begins to reconsider the importance of having the learning outcomes to hand:

*“... level six critical evaluation of practice specific formulary, somewhere is the learning outcomes. Got to, got to find these. Oh here they are ...”.*

The above excerpts present a case for interviewees having sight of the learning outcomes to ensure satisfactory evaluation of a written assessment artefact. Conversely this can also pose a problem when there are a large number of learning outcomes which are difficult to select evidence that demonstrates achievement within the context of the assessment task set. The interviewees who were unable to access the learning outcomes for the assessment task appeared to experience greater difficulty in evaluating the assessment artefact.

#### 6.3.1.e MARKING AND GRADING ENVIRONMENT

Interestingly only one interviewee explored the issue of environment, ‘Maz’ is quite honest when exploring the potential issues:

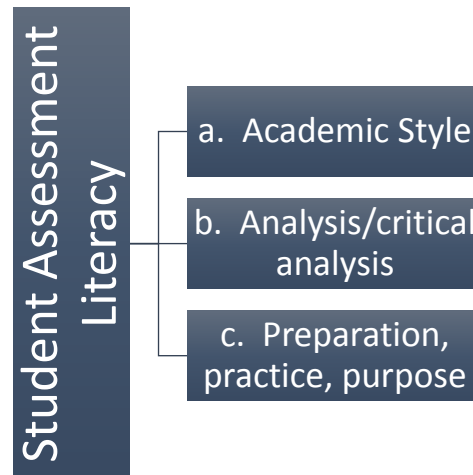
*“... do you sit here [**in a shared office**] and do it where you are likely to get disturbed and we all have this dilemma actually do you say I am going to stay at home and I am going to do them at home but I think it is it is difficult when you are [**here and you get**] distracted so you might be half way through and the phone rings or something ...”.*

This sub theme was not significant for all but two PA interviews as the others all took place in small study rooms. One PA interview was conducted in their own home whilst the other was completed in an empty but shared office (‘Maz’).

Overall this major theme explores a number of environmental elements which have an impact upon the completion of marking and grading by the interviewees. Each of these elements exist independently of the individual

lecturer interacting with the assessment artefact. The existence of one or more of these named elements has the potential to affect the processes of marking and grading of written assessment artefacts.

### 6.3.2 STUDENT ASSESSMENT LITERACY



**Figure 13 PA Major Theme**

This major theme explores the sub themes identified above, which in one way or another play a role in the evaluation of a written assessment artefact as identified by interviewees.

#### 6.3.2.a ACADEMIC STYLE

This theme explores how interviewees consider the structure and language of the written assessment artefact very early in the marking and grading process gives a particular impression of the value of that artefact. This is very evident in an extract from ‘Maz’:

*“...This particular one what I have picked up even from just looking at the first couple of pages is her writing style isn’t as I would really like it but I am quite critical in a way. Sometimes people write very colloquially they don’t structure their sentences terribly well ...”.*

This perspective is continued when ‘Maz’ details further elements of the written assessment artefact which present a particular impression:

*“... I guess for me as well the ease of reading if it is well organised well- structured if it reads well makes it an awful lot easier to follow*

*doesn't it and kind of guide your reader through things which I don't think they do very well in general ...”.*

Further issues about structure and choice of presentation of information within the assessment artefact itself are highlighted by 'Shirley' who says:

*“... because of the appendix it probably loses its flow ... this is a bit irritating having to keep going to these appendices but I kind of think that's her choice, the thing is by the time I get here I forget what appendix I'm looking for ...”.*

'Toni' highlights how having an assessment artefact which is easy to read and understand has an impact upon the subsequent papers they come into contact with:

*“...if it's the first one I've marked it's easy to mark if you mark a good one first and then in its very easy to compare and contrast isn't it ...”.*

Conversely where there are more issues identified with the written assessment artefact 'Toni' expresses the difficulty that this poses:

*“... so if you mark a bad one first, you don't know it's a bad one there's a lot of checking backwards and forwards is it good or is it bad until you've sort of identified a marker post in your mind; this is what good looks like, this is what bad looks like and this is what in the middle looks like if you've **[student]** not done too much of this but you've done some of that ...”.*

For Sam the way an artefact is structured and the language it uses are a barometer which influences their perception of the assessment artefact itself in relation to the final summative assessment:

*“...there was no surprise in the mark that the paper achieved. If it had come out at 40 then I think that would have been too low, reading it I would've expected it to come out at 60 or 70 most of the things that needed to be there were ... and guess the writing style itself was good but the actual flow from one section to another didn't necessarily logically follow so yet it seemed to come out about right ...”.*

Style, structure, use of academic language influences the perspective taken on the written assessment artefact. Interviewees disclosed how these three

elements lead to an increase or decrease in the ease of interaction with the assessment artefact itself.

### 6.3.2.b ANALYSIS/CRITICAL ANALYSIS

Learning outcomes which include the words analysis or critical analysis are commonly expressed and require a response being evident in the written assessment artefacts. These are important elements to identify but responses could be multifarious. It appears students experience difficulty in demonstrating that they can achieve the process of analysis. For instance 'Ali' identifies within the artefact they are evaluating the following:

*"...Descriptive, descriptive, descriptive, interesting descriptive ..."*

Interviewees demonstrated that they had different expectations for different levels of assessment; again 'Ali' provides evidence of this:

*"... now interestingly this is level six so the level of synthesis I would be expecting wouldn't be as great as for a level seven. But I would still like to see a little bit of kind of critical kind of about why ..."*

Interviewee 'Jo' gave an example of what analysis meant in the assessment artefact they were reviewing:

*"... the student needed to go one step further and after identifying what biases existed think about how that affected or could affect the research and the results and suggest what else could be done to negate those biases. It is probably to do with having a better understanding of research processes so that they could present a reasoned or rational suggestions for changing practice ..."*

This example shows how doing more than including a description of the issue was required to demonstrate the theme of analysis or critical analysis in written assessment artefacts.

In a number of cases, interviewees were left wondering whether there was evidence of analysis or critical analysis and what this should look like and how to communicate this as feedback. As 'Ali' puts it:

*"...you know kind of sort of well into the first page, here I would expect to see some level of **[some]** kind of analysis going on ..."*

Even when the interviewee identifies that the topic is not within their specialist knowledge they can see opportunities for analysis which are not taken by students and interaction with the subject remains largely descriptive.

#### 6.3.2.c PREPARATION, PRACTICE, PURPOSE

This sub theme highlights three components which have implications for the assessment artefact. For instance one interviewee identifies that poor academic writing has an impact on the final mark or grade to be achieved.

'Ali' suggests:

*"... a day on academic writing doing a marking workshop ...[using] previous ... failed, borderline, good and very good and give them a criteria and get them to work in groups ... before they submit their essays..."*

Undertaking such an exercise would develop student assessment literacy, by highlighting expectations and developing understanding of the application of marking and grading criteria.

The next issue to be tackled is the use of direct quotations, which students with poor academic skills use frequently. As 'Ali' highlights:

*"...there are sometimes where a statement has been made so beautifully there is no better way of saying it and at those times or if you are defining something umh I think that direct quotes are great..."*

Unfortunately direct quotations do not provide evidence that what has been written has been understood and can be utilised or reapplied in a different context.

The final component to be addressed under this sub theme is related to ensuring students understand the purpose of receiving a mark or grade, to provide evidence of achievement or otherwise. One interviewee explores the approach of students who receive a mark which does not reflect a pass mark for the assessment. 'Maz' suggests that:

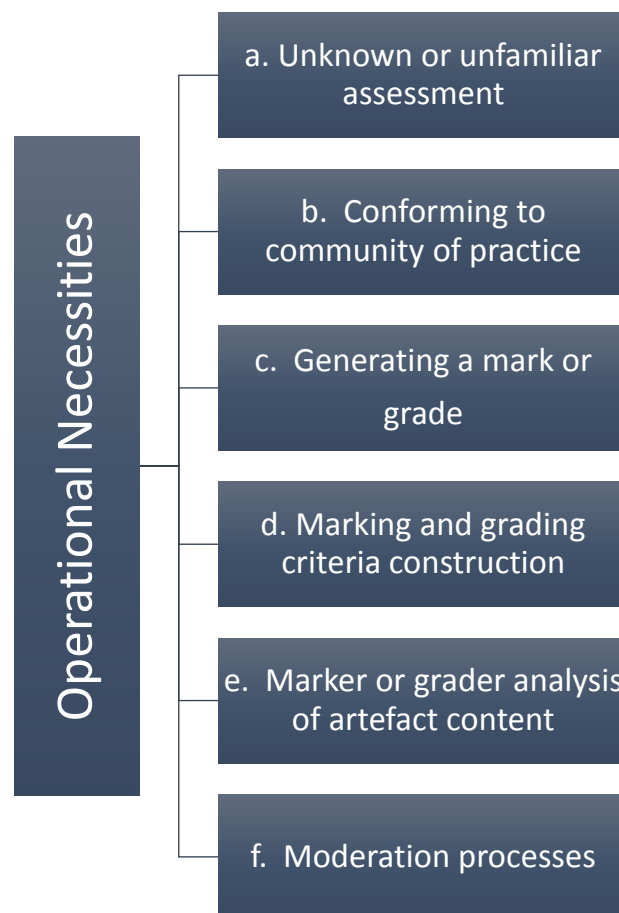
*"... if you give them something back at 38 they will think well I only need two marks to get it right ..."*



The potential here is to communicate an inaccurate message, that there are limited changes to be made to achieve a pass mark of 40 or greater for the assessment task. Developing students' assessment literacy would facilitate a more accurate evaluation of the improvements necessary to move their marks by two percent or greater.

This sub theme briefly examines issues which are related to students' understanding of the assessment task and artefact, but potentially impact upon marking and grading completed by lecturers responsible for evaluating assessment artefacts.

### 6.3.3 OPERATIONAL NECESSITIES



**Figure 14 PA Major Theme**

This major theme is concerned with physical, contextual or procedural elements which occur prior to or simultaneous to crafting a final mark or grade for award to written assessment artefacts in this research study. The

sub themes all appear to be aspects to either be aware of or to complete as a lecturer executing their role in marking and grading.

#### 6.3.3.a UNKNOWN OR UNFAMILIAR ASSESSMENT

Interviewees disclosed their sense of anxiety when confronted with a topic area of which they have little or no knowledge and yet are expected to provide a mark or grade whilst evaluating an assessment artefact. 'Jae' sums this up:

*"... slightly apprehensive because I have never worked in ... I am thinking straight away I perhaps need to look at some of their references ... because they could in theory be saying anything where if it were on a subject [I knew] ... I would have more confidence I perhaps wouldn't be so apprehensive ..."*

'Shirley' identifies that limited knowledge about a topic area could be more to do with:

*"...that it might be an up-and-coming treatment and I'm just old and do realise that ..."*

This interviewee is acknowledging that continuous changes in practice render their knowledge as less than contemporary.

Anxiety is present for interviewees when they are completing marking and grading for where they are unfamiliar with the module content or structure for the assessment artefact, for 'Toni' this proved a problem:

*"...I had a look yesterday and tried to get the module descriptor but couldn't ..."*

Unavailability of detailed guidance resulted in further confusion for 'Toni' when exposed to the written assessment artefact:

*"... I'm looking through the paper I also was confused by the fact have got two submissions ... because the word count is limited to 2000 words but the module leader has got around it by saying that can submit a separate introduction ... seems a bit strange ..."*

In contrast 'Ali' identifies how important it has been for them to have the:

*“... the learning outcomes are just in my head ... it might be very different if it wasn't a module that I know inside out That influences it as well ...”.*

For this interviewee knowing the assessment requirements well became important in being able to decide on the success or otherwise of the artefact at achieving those requirements. Knowing the module learning outcomes and the assessment artefact task, facilitated this interviewee in awarding a mark or grade which denoted a failure to achieve the required level.

Whilst 'Maz' expressed their anxiety due to poor knowledge of the assessment task:

*“... So I am a marker ... I am not really part of, I have very little input into the actual module itself. We are allocated about eleven or twelve students to mark for ...”.*

This position is contrary to that of the previous interviewee 'Ali', who feels more confident in their task because of their underpinning knowledge of the assessment artefact requirements.

Unease for these interviewees took different forms, still however having the potential to impact upon the evaluation and awards of marks or grades for an assessment artefact.

#### 6.3.3.b CONFORMING TO COMMUNITY OF PRACTICE

The topic selected for this sub theme appeared to be broached by two interviewees. 'Maz' first raised the issue relating to unsuccessful assessment artefacts and a comment that was made to them:

*“... I think I failed something and I gave it about thirty eight and somebody said no if you are going to fail it give it 35 ...”.*

This approach appears to be relevant to an issue explored in a previous sub theme (6.3.2.c) and the perception of the student.

A different approach is taken by 'Toni' which still demonstrates engagement with the community of marking and grading practice within their department, they:

*“... mark this on my instincts and then I'll go and look at one that's been given a mark similar to what I've given to see if it's comparable so that will be my checking process and then I would get the second marker to definitely consider one of the two that I'd marked ...”.*

This sub theme illustrates a concern by either interviewees or their colleagues to ensure that marks and grades awarded are acceptable enough to be awarded from their department.

#### 6.3.3.c GENERATING A MARK OR GRADE

This sub theme is explored through extracts from five interviewees and involves the identification of how they each were able to produce the mark or grade to be awarded for the assessment artefact.

For instance this interviewee is reviewing what has been written and is talking out loud identifying this, at the same time as talking of the decision they made:

*“...Although it is not bad ... it is kind of clunky. It is not unprofessional. But I wouldn't have said it was very good. I would say it was good use of professional and appropriate language ...”.*

This is 'Ali' as they talk about what has been included in the artefact and their estimation of the level of achievement, finding where it fits on the marking and grading criteria. That being said, once each category on the marking and grading criteria has been decided upon, it is still left to the marker or grader to decide on what final mark or grade will be awarded. This point is well illustrated, again by 'Ali':

*“...there's five different items in that box, do they all have similar weighting when you are thinking about your overall mark ... you can present an argument ... compared to whether they have managed to analyse it I think it's much more important ... So even though he has got a 3 in that column I would put him at the lower end of this column. Probably 14 again that's a big range but middle of that column probably which would take us to sort of 11 ...”.*

In spite of making a decision regarding what was felt to be the most important elements to detect in the assessment artefact, further decision making is required, as to how much or how little to award. Whilst 'Jae' describes another approach for the awarding of marks or grades within separate criteria and for each element within those criteria, identifying there are two approaches to this:

*"... start at a hundred and ... knocking marks off for everything they have done wrong or you start at zero and are giving marks for and adding it up you know ..."*

There is the potential for these two differing starting points to end with a dissimilar mark or grade being awarded for the same assessment artefact.

From reviewing the transcript of 'Maz' it can be seen that they are comparing elements of achievement of certain criteria within an individual assessment artefact. This is evidenced by the following excerpts from 'Maz's' interview:

*"... I would say she is somewhere around here with knowledge ... it is acceptable but ... not wonderful yeah. So ok she has got some evidence ... not terribly well substantiated. She's got some analysis going on ... but again around about the same point. Originality is always one that is I don't know really. Somewhere sitting on the line I would think. She has got an awareness ..."*

This conversation with the assessment artefact continues until the interviewee has reviewed all the required elements on the marking and grading criteria and settles upon an evaluation, 'Maz' states:

*"... I would put her somewhere between 40 and 50 again I would think. It is tending to me to feel there ..."*

This approach is evident of the interviewee using criterion referencing to structure the mark or grade to be awarded. A careful checking process with the artefact being compared to each element on the marking and grading criteria facilitates decision making, though the above excerpt remains a 'ball park figure' rather than an end product.

In contrast 'Shirley' is undertaking a different process in order to support her decision or judgements on assessment artefacts that they have been marking and grading:

*"...because this is the last one I've marked and then you start thinking of one compared to the other and I would say compared to the ones that I've marked that this would come out with the lowest score I do think that it's better than that ..."*

The process described is of comparing achievement of one artefact against another, rather than simply relying on the criteria provided for marking and grading to decide on the final award.

Marker or grader bias is another aspect that comes into play during the evaluation of a written assessment artefact. This point is illustrated by 'Shirley' who wonders:

*"... have been a bit mean, I think the section that maybe I could go up on here is her relevant analysis ..."*

The interviewee continues to explore their original decisions:

*"...she has done a table and has ... had a good look at the papers ... she did explain all the outcome measures I just didn't like where she put them that doesn't mean that it was actually bad she still had them ..."*

From this, it can be seen that because 'Shirley' disapproved with the way the work was presented, they did not read or see how well the content fitted with the task set and therefore did not initially award a commensurate mark or grade:

*"...bring that up a bit so she's come out with 65 that's given a big jump but would sit more comfortably with me that sounds better this ..."*

The complexity of generating the final mark or grade is illustrated by 'Jae' who even at the point of commencing to draft the written feedback is still unsure of the value of the assessment artefact:

*"... I have not really made up my mind yet, sorry, on grade yet. Thinking. Don't think it is a first. But I think it is in the sixties somewhere but I may I can be swayed ... So just looking at the*

*marking ... I am kind of borderline at the moment is it a first or a two one.... I know it is not seventy in the intellectual skills. So I just need to be aware of the words I use here [in the feedback] ...”.*

This difficulty with decision making is not unique to papers which are on the cusp of achieving a higher grade. This excerpt from ‘Ali’ is reflective of the dilemma for lecturers marking and grading:

*“... Oh heck 39 so we have a border line assignment here ... I think 39 is very tricky ...”*

After some consideration ‘Ali’ decides:

*“... leave him teetering at 39 because it is a full portfolio and I can come back to this one at a later point so my gut feeling is that we could tweak things to get him into a pass ... a low pass, but I am going ... look at how the rest of the portfolio informs on it ...”.*

This sub theme presents a variety of considerations which impact upon the producing of a final mark for an assessment artefact and demonstrates some of the difficulties that can arise.

#### 6.3.3.d MARKER OR GRADER ANALYSIS OF THE ARTEFACT CONTENT

Two interviewees had excerpts coded to this sub theme. There were extensive examples from ‘Ali’, as there were three pieces of work which made up the complete assessment task; the protocol analysis interview was completed in an interview of over 2 hours long. Some of the comments made by this interviewee were related to them exploring the assessment artefact and others were more focused on the content. As an example ‘Ali’ explores the artefact and the authors approach to reflection, relaying that some assessment participants would be able to follow a reflective framework:

*“...but then as they become more used to it actually reflecting continuously throughout ... this feels very much like ...[the student saying] now I am going to do a bit of reflection cos that is what we are supposed to do ...”.*

This suggests an approach to task completion by closely following the task requirements and then being constrained by these, demonstrating limited individuality, but is an expected attribute in a reflection.

This interviewee then goes on to give examples of their application of the marking and grading criteria, by considering each element and then comparing sections from the assessment artefact:

*“... application of knowledge I think this is interesting cos ... the bit in this assignment where he has applying his knowledge ... where he talks about ... calculate the dose ... must consider fat distribution because of the nature of the drug ... a really neat example of him applying his knowledge and of relating it to the case study so I think that is quite good ...”.*

A similar process is relayed by ‘Shirley’:

*“... so she's mentioned ethical issues but it's all just a bit superficial ... if I don't understand what she's talking about I don't know whether she does but ... it's not coming across ...”.*

These final two examples have been related to exploring the content of the assessment artefact, in light of the criteria against which it is to be judged.

#### 6.3.3.e MARKING AND GRADING CRITERIA CONSTRUCTION

This sub theme identifies the marking and grading criteria in use for the assessment artefacts being evaluated. Each interviewee took a different perspective when they talked about the criteria which were available to them.

‘Jae’ described the criteria they had access to, it is set out across an A4 sheet of paper:

*“... there is very good, excellent, good whatever and then there is a description saying what that is ... then basically you read the assignment and you obviously use the learning outcomes as well but you look at it and you basically say which one of those this assignment fits in ...”.*



With this criteria, the lecturer has to take a global perspective of the assessment artefact in order to identify the mark or grade to be awarded.

'Jae' then continues:

*"... I struggle with that because I think that elements of it, knowledge and understanding, could potentially be excellent to very good but then transferrable skills may be very poor and then where do you fit it on this it is very much then your own interpretation how you, how much you weight it and I know a lot of lecturers really like that style but I don't find it very useful ..."*

This interviewee reveals the difficulty they face with a set of criteria which appears to offer a great deal of flexibility. However, this interviewee views this manipulability as detrimental to the outcome of their assessment artefact evaluation due to their lack of marking and grading experience.

For 'Maz' the marking and grading criteria to be used is put together from available guidelines, but being new to this particular module, they decided to review this later:

*"...we have a marking criteria we copy we have a guideline and then we have some weighting and structure on the marking ... I will probably look at that later on about the way that the marks are allocated ..."*

The interviewee made comment on what was done well, and what needed further work, this information was then used to construct the final feedback and marks or grades to be awarded for the assessment artefact.

The interviewee 'Toni' was reviewing an assessment artefact for a module which they had had no previous contact. On interacting with the marking and grading criteria for that work they were surprised:

*"... this marking grid, it's not a marking grid that I've seen before, it's like the difference must be specific to this module or postgrad case study kind of things ..."*

From this excerpt it can be assumed that 'Toni' had not familiarised themselves with the marking and grading criteria for the assessment artefact, submitted for a level they did not commonly complete evaluations at.

A further aspect, not as yet considered here is the impact of how the criteria is constructed i.e. how it is put together and its influences on marking and grading, as 'Maz' highlights:

*"... I feel that the words within this ... I feel that the wording tends for me to pull it downwards ... we have new marking criteria coming for the summer ... I think the wording within the new ones is more positive ..."*

This interviewee is awaiting the introduction of new criteria which they hope may improve marks or grades awarded for an assessment artefact, due to an emphasis on positivity rather than identification of omissions.

#### 6.3.3.f MODERATION PROCESSES

A process of ensuring the legitimacy of marks or grades awarded to assessment artefacts following evaluation by the interviewees is examined here. The moderation process takes many forms, one which was totally unexpected, when I seemed to become the barometer or gauge during the protocol analysis process for 'Ali':

*"... Warfarin is used to treat several conditions including DVT, interesting. Are you with me on failing this one?"*

From the perspective of this interviewee, they had had another academic present when reviewing all 3 assessment artefacts which contributed to the final requirements for the post graduate programme they were evaluating. Having another academic present seemed to provide some reassurance to 'Ali', who although not new to marking and grading, was conscious of the ramifications of awarding a fail grade. In line with institutional policy this assessment artefact would still be subject to second consideration.

'Jae' held a perspective that there was too much second marking and grading:

*"... that makes me feel you **[the institution /department]** are not confident in the first marking ... then you get these big variations in grades and you know the two first and second markers getting anxious that their marks are so varied ... "*

One explanation for the variation in evaluation of assessment artefacts is the different focus taken for marking and grading by individual assessors.

'Toni' has an interesting perspective on this point:

*"...if the first time I was marking this I would sit down and discuss with the module **[leader]** and even then one person's gold is another person's average ..."*

Individual academics focus on different elements within the assessment artefact, for instance writing style, referencing, inclusion of research, or content. The first marker will have had their own focus along with taking account of learning outcomes and any specific instructions for composition of the artefact. Second marking and grading entails a detailed review, with the lecturer making comments about the artefact's merits or disadvantages, individual foci will come into play here. A large variation in marks is less likely to occur with a process of second consideration, where a more superficial review of the assessment artefact occurs but does include a review of the qualitative feedback.

This difference between second consideration and second marking is highlighted by 'Maz' as they explore how they were supported in the marking and grading aspects of their role:

*"... what we have within the department here is second consideration. So it is considered rather than marked in huge detail. But kind of your first year while you have got an academic mentor then people did properly kind of second mark them for me ..."*

'Maz' indicates that second consideration can be an invaluable tool:

*"... the thing for me ... is if I am not sure, even if I think that it is a pass, but I am still not quite sure where to put it then that is where the second consideration **[comes in]** and we sit down and as I say I have not been too far off the mark ..."*

However, this process is not fool proof as 'Maz' highlights the second consideration process does not necessarily promote greater objectivity, but could result in continued subjectivity:

*“... well the student had missed the point completely of the assignment so that was one thing. Plus even if you read it to try and give it some academic consideration in the form it was in, it was still way way off. Well what do you do with that? So somebody said give it twenty. Somebody else said well give it thirty ...”.*

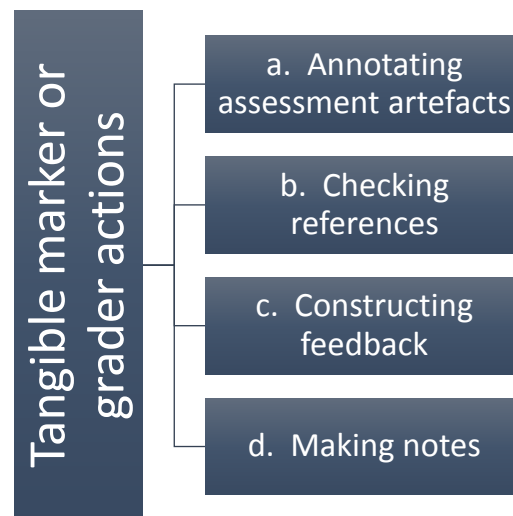
It is in making such difficult decisions that the process of moderation is most required. This could be second consideration but at times may need to be second or even third marking to reach an agreement on what the artefact should be awarded.

Finally a sentiment that is well articulated by ‘Toni’ is that uncertainty:

*“... I feel, that if no one else checks it what about that student as I've marked it, is it right, is it wrong, have I been overly harsh, but looking at that other 40% one, they've [**another lecturer**] slated it. Let's look at that 80% one, they have a nice concise intro, got the terminology in straight away, references, straightaway they've looked at the right thing ...”.*

A number of strategies were operationalised by interviewees to facilitate moderation and increase confidence in the accuracy of marks or grades awarded to assessment artefacts. Not all interviewees broached the subject of moderation. Generally those interviewees who felt unsure about the mark or grade they were awarding or those who had evaluated an assessment artefact to which they had awarded a fail raised the subject of second consideration.

### 6.3.4 TANGIBLE MARKER OR GRADER ACTIONS



**Figure 15 Major Theme**

This major theme was constructed through identifying the physical rather than entirely cognitive aspects undertaken by interviewees completing marking and grading. Extracts from interviewees will be used to explore and illuminate the individual sub themes identified above.

#### 6.3.4.a ANNOTATING ASSESSMENT ARTEFACTS

Four interviewees explored the subject of annotating assessment artefacts including its merits and disadvantages. Extracts from the interviewees highlight how interviewees utilised annotations on assessment artefacts.

Whilst observing ‘Sam’ they disclosed that they were making notes on the assessment artefact but that this was not routinely sent back to the student:

*“... I don't think they're getting it back specifically but do tend to put at the bottom of the rubric that there is an annotated version available for them to look at ...”.*

For ‘Sam’ the purpose of placing annotations on the assessment artefact was to facilitate the development of feedback. To ease this process they identify:

*“... if there are key points I think I need to make, I have got another word document open that I'll then start building some of that in so I can just copy and paste that into the rubric as I'm going along ...”.*

From this excerpt it is clear that this interviewee has a process, the first stage of which includes making comments on the assessment artefact. The next stage is making a note of key points or issues on a separate word document to then, transfer it to the marking and grading rubric. Annotations act as a short summary or reminder which can then be used to construct feedback and inform the award of marks or grades when reviewing criteria on the rubric.

One interviewee recounts their experience of trying to use annotations on a written assessment artefact to support students. 'Jae' uploaded annotated documents to the virtual learning environment for the assessment artefacts that they had evaluated:

*"...the module leader was saying rightly or wrongly everyone should do it or no one should do it. And also a couple of students complained ... so I have stopped doing that which I am not convinced is right ..."*

In the spirit of equality the practice of providing annotated scripts only to a proportion of the module participants is discontinued. 'Jae' also commented that there was a potential that students, without further explanation:

*"... would perhaps interpret something off my annotating differently to the way I wanted .."*

However, 'Jae' still felt that despite that possibility they:

*"... can see the pros and cons of it but I think .... it is nice to see that the person marking it has clearly read the whole thing cos I am commenting on and highlighting you know the points that ... are good and bad points that are strong ..."*

As well as being available for the student, 'Jae' finds the annotations useful:

*"... as a reference point so when I am giving feedback, so that is how I do it ..."*

Other interviewees discussed how they used annotations on assessment artefacts to ensure that authors received feedback which is more helpful, 'Jo' states:

*"... I'm wanting to make it more specific to the paper I think in some ways they will hopefully get that from reading the annotated script,*

*that's the purpose of always liking to do an annotated script in some ways it's more useful than the small degree of comments here [on the rubric] ...”.*

This interviewee always ensures that these annotated artefacts are returned to the student:

*“...load the annotated scripts up to, where the rubric or the feedback sheet is. Yes I will upload it I always used to send it directly back to them as an email attachment, but now it gets uploaded so I just uploaded onto the system ...”*

It seems that inequities exist between departments and between institutions in the use and utility of annotations and their return as feedback to students.

One interviewee considered that putting comments on assessment artefacts was a way of developing their skills; they greatly appreciated feedback they received from assessment students:

*“...the feedback they give me is that yes, thank you that's what we need to know ...”*

From the extracts above it is surmised that annotations are developed as lecturers review assessment artefacts and are used primarily to summarise attainment or otherwise and to assist in constructing feedback for marking and grading rubrics. Whether or not students gain access to this resource appears to be a local matter.

#### 6.3.4.b CHECKING REFERENCES

The review and inspection of references including citations and the reference list is discussed by 8 of the interviewees, extracts from 7 are presented to illustrate differing levels of attention. ‘Jackie’ appears to take a more global perspective:

*“... as I am looking through it generally I am looking at references. Occasionally I will flip to the reference page to see that they are referencing it ...”.*

In contrast Jae:

*“... their first reference it’s a minor error not putting a title in italics so ... I’ll perhaps maybe going to focus on have they referenced correctly throughout now cos the first one was a minor error ...”.*

Other interviewees were much clearer about their processes, ‘Ali’ provides details:

*“... So what I usually do when I start to mark is the very first thing is just have a very quick whizz through the referencing before I read the thing at all. It is kind of administrative ...”.*

On that first peruse through the references, both cited and on the reference list ‘Ali’ soon identifies:

*“...I can instantly see he has got dates missing from the references. From all of the references in fact ... making me think back to ... the guidance they were given about academic writing ... seven references for a two thousand words case study intuitively feels like not really very many given that one of them is a summary of product characteristics ... as is another one, and another is a BNF [**British National Formulary**] so nearly half of them are just descriptions of drugs ...”.*

As this interviewee reviews the assessment artefact and identifies there are few sources of literature used to support the content with 3 of the references cited being pre-printed material giving drug information. ‘Ali’ appeared to be disappointed with the amount and type of literature used to support a reflective case study along with the citation errors given the guidelines provided at the module outset.

‘Jo’ is another interviewee who carefully checks each of the citations present in an assessment artefact, they say:

*“... I tend to do this first, I like to make sure that they are all as they should be I check ... that they match with the references on the reference list. So if they’ve got et al. there should be at least 3 authors and that they have all the correct information ...”.*

This interviewee then talks of the difficulties in carrying out this process and the change in their approach:



*“... there is a lot of 'uping and toing and frowning' when I'm starting going through this. This is quite a laborious process going through ... but its the only way I've managed. I used to look at the references when reading through the paper but ... it broke me off too much ...”*

It appears that before reviewing the assessment artefact for achievement of the learning outcomes, interviewees spend a portion of time checking through references in both the list provided and citations in the text. The majority of interviewees undertook a detailed examination of references including checking for consistency, appropriate use of et al., use of grammar and use of a letter to denote the differences between the same author and year but a different text.

An example of how much effort interviewees go to, to identify literature on the reference list presented with an assessment artefact is illustrated by 'Jae':

*“... now what I do I'll look at the reference list ... count them ... and then what I do is I just pick out completely random references from the study just to see so Kirby and Blackburn 1987 are cited in Cook so let's look at Cook ... so I know that that is a book so I am not going to be able to get that to look at ... I really want the article to try to see if she has referenced the Cochrane Review which I don't think she did. Hang on Cochrane she didn't reference it properly so I just type that in [Google/Google Scholar] and I'll see if that comes right ...”.*

The time available to undertake such detailed exploration of the references submitted with an assessment artefact will be influenced by interviewee workload, knowledge of the module assessment and or content of the assessment artefact. 'Maz' concurs with this perspective:

*“... I am looking as well at where she is getting the references from. This is something I know quite well where has she been for the supporting kind of evidence it is important ...”.*

'Toni' is somewhat at odds with the rest of the interviewee cohort who:

*“... still read the reference list carefully there's quite a few errors ... but ... well I know how hard referencing is so they've had a go and*

*then been consistent with themselves and I can make sense of it and I can find the document then I'm not penalising them for it ...”.*

The interviewees demonstrate an inconsistent approach to the identification and checking of references cited in the text or on the reference list. This inconsistent approach continues with how or whether interviewees identify if a legitimate reference has been used. There are difficulties in negotiating the assessment artefact, to review individual references and this appears to hinder the process. One interviewee perceives that there is inherent difficulty in completing citations correctly and because of this perspective they are reticent in penalising the author of an artefact where the correct format has not been used.

#### 6.3.4.c CONSTRUCTING FEEDBACK

This sub theme presents a small selection of extracts from the data around the theme of constructing feedback. This sub theme is explored as it provides a window into one of the features of marking and grading a written assessment artefact. Two of the interviewees consider how the student will respond to feedback provided. For 'Ali' they reflect on their own experience:

*“... I remember as a student you just want to scan it and you don't want 'on page 19 you refer to such and such' [or] 'I suggest you read around such and such' because actually you are thinking ... I just want to know if I have passed. So I do think we imagine that quantity means quality and I don't think it does ...”.*

The suggestion here is that a minimum amount of feedback is all that is required, with little need to identify where the artefact could have been improved. When there has been a successful assessment event, feedback forward (as in what could be done in the future) from this interviewee's perspective is of limited use, as its value may remain unappreciated.

Whilst another interviewee 'Jae', does consider the student but takes a different perspective:

*“... so I what I do on my feedback often rather than saying you showed such and such I will just perhaps copy and paste an element*

*of the assignment and say this is good and highlight it.... to show that this is what you have done and should try and do more of it ... something wrong then show it, copy it directly and highlight it and try and show ... cos you don't know what they are thinking of the feedback do you ...”.*

Both interviewees are considering how the student will react to the feedback. ‘Ali’ appears to think they will concentrate only on the mark or grade awarded, rather than any information on what could be improved. The extract from ‘Jae’ demonstrates the opposite perspective, as they carefully craft feedback to illustrate the areas of strength and those requiring development within the artefact.

The interviewee ‘Jo’ discusses their rationale for providing feedback on an assessment artefact, they explore how they would advise the student on how to improve their work within the specified word count:

*“... one of the things I'm always aware of when doing feedback is ... a tendency to say you need to do this, you need to do that ... make sure that I highlight the things that possibly were not necessary and could have done without ...”.*

Considering the impact of feedback on students appears to be the theme of the preceding interviewees, whilst the following interviewees explore issues with themselves as the constructors of feedback artefacts.

‘Jackie’ discloses how they struggle with compiling feedback on assessment artefacts:

*“... [it's] my weakest area. It is something I need to develop ...”.*

For ‘Shirley’, the issue is slightly different, but potentially has the same results and outcome as for ‘Jackie’, difficulty in constructing feedback:

*“... sometimes I look at it I'm not just quite sure which bits to put it in say analysis and argument...”.*

The final interviewee broached the idea of the differences between feedback on draft work and on the final submitted assessment artefact, and being

unsure how much to do. This extract from 'Maz' presents these anxieties clearly:

*"... it is the difference between having just ... looked at a draft of work where you have thrown lots and lots of comments on and ... what you write on a final marked essay I find quite confusing ..."*

It appears that 'Maz' has identified the disparate purposes for feedback but has concerns:

*"... I am not sure and I would love to know how much they really read what you say about what the comments are on the final piece. If they have passed do they really look at it ..."*

They consider that their more experienced colleagues provide limited feedback or less volume on the final piece, although they concede:

*"... It's a very varied thing between people about what you put and again it is one of those things where there is no rule and there is no real guidance ..."*

This sub theme has explored interviewees' perceptions of crafting responses to written assessment artefacts. All of the interviewees raised concerns which took two forms, firstly concerning the impact of feedback on students and secondly individual issues regarding interviewees' perceptions of their ability to develop quality feedback.

#### 6.3.4.d MAKING NOTES

During protocol analysis interviews it was evident that interviewees made notes when reviewing assessment artefacts. In the main these were hand written notes, as 'Maz' highlights:

*"... I have a piece of paper and I do make some rough notes as I am going through ..."*

This practice is supported by 'Jackie' who talks of writing notes to themselves. The type of notes are generally related to the assessment artefact under evaluation, 'Jo' for instance is:

*"... making written comments about grammatical errors ... picked out one of the key areas which was done badly in the paper ..."*

The notes act as a memory jogger for 'Maz':

*"... as well if there is something kind of particular it just helps me ...".*

Whereas rather than making handwritten notes 'Sam' relies on using a word document to write observations and comments:

*"...if there are key points I think I need to make have got another word document open that I'll then start building ..."*

The majority of interviewees made notes, often via pen and paper to remind themselves of aspects they wanted to comment upon or review further on in the artefact itself.

This sub theme briefly touches upon an aspect of practice that is useful to record. Marking and grading is normally a solitary process, using protocol analysis provided an insight revealing a common act, only known by those who evaluate assessment artefacts.

### 6.3.5 IMPLICIT KNOWLEDGE OR ACTIONS

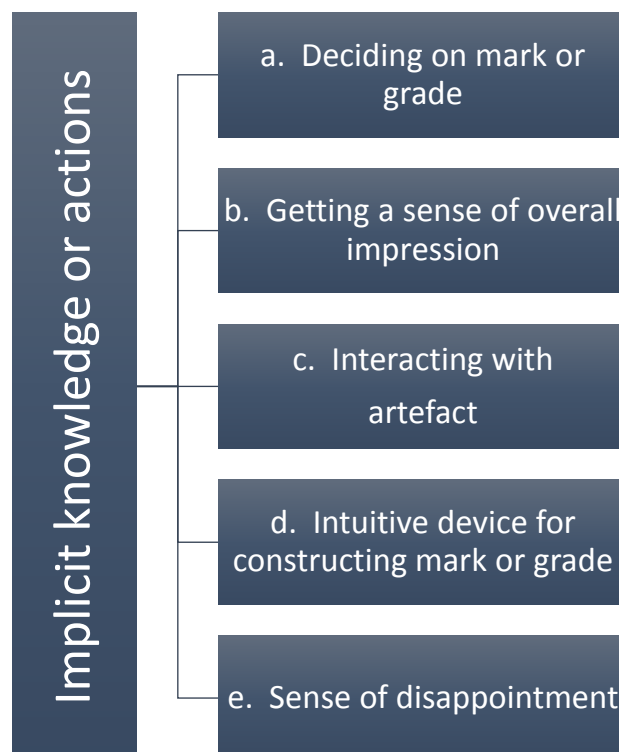


Figure 16 PA Major Theme

This major theme is concerned with exploring all the actions which were led by cognitive drivers and gives the opportunity to review these processes by exploring extracts from interviewees. As the researcher, I have given a name to the processes and cognitive actions I identified within interview transcripts, allowing the information to be divided up into a number of sub themes.

#### 6.3.5.a DECIDING ON THE MARK OR GRADE

Interviewees in a number of ways talked of the difficulty of coming to a decision on the final mark or grade to be awarded to the assessment artefact. For instance 'Jo' confesses:

*"... I always have problems with putting someone in a category, especially when there is not a clear border ... you've got to gauge ... one of the difficulties I have is ... judging ... you kind of give them the benefit and move them up, but if you've done that a few times, you'll probably pick the lower one on the next. It's not an ideal system, so yeh I kind of find that difficult ..."*

This excerpt highlights that providing a mark or grade is a complex interaction of balancing marker or grader perception of the artefact with the criteria available to describe the achievement. This is evident when 'Jo' talks of how at times they may be lenient ("*give them the benefit*"), but in order to achieve a balance at times may select a category or criteria that may be a little lower in describing achievement. Ultimately this is in an effort to achieve a balanced judgement as to the value of the assessment artefact.

'Ali' made an interesting comment regarding an assessment artefact which they are reviewing where the author is a post graduate student:

*"... in terms of academic work he is where a first year XX might be. So my expectations of a XXY's portfolio are different even though the marking criteria are the same because they are about academic work the depth of reflection I'd expect from a XX is different ..."*

'Ali' is voicing the opinion that despite being assessed at the same level for a post graduate programme, they would have different expectations of participants. The rationale for this is differing ability to write an in-depth

reflective account which would have held dissimilar levels of importance in the different profession's under graduate education. This approach will have an impact on 'Ali's' decision making with regard to what mark or grade to award. Interestingly another lecturer without 'Ali's' implicit knowledge may view the assessment artefacts in another light and be less inclined to make allowances corresponding to those of 'Ali', thereby resulting in a different mark or grade being awarded.

To facilitate 'Sam' in making decisions about what marks or grades to award they access:

*“... [a] high and low from last year's marking so that I could look through to see what sort of things the student was been asked to do. Some prepping in relation to thinking about what I would expect from high-level students and what I would expect ... from students who didn't achieve a higher mark ... would go through a paper myself just so that I could look at what things are needed ...”.*

This interviewee reviews this information prior to commencing the evaluation of assessment artefacts in the current academic years' submission. For some this would seem to be an onerous undertaking, increasing their workload. Whereas for others, such as 'Sam', it is a necessary component in their marking and grading process, increasing their confidence in the decisions they make.

For 'Maz' they continue to struggle with identifying exactly where to place their assessment by using the marking and grading rubric, whilst it has become a little easier:

*“... think it is probably one of my biggest dilemmas ... I think the actual finding the matrix [rubric] and kind of placing it has got easier for me but actually me putting my finger on that final mark is quite difficult ...”.*

How that final decision is reached remains elusive to 'Maz' who continues to be unsure and have limited confidence in the final mark or grade to be awarded. They talk about discussions with colleagues to gain support in the decisions.

'Jo' explains a little more about how they decided on the mark or grade for the assessment artefact used in protocol analysis:

*"... 51, which thinks seems reasonable and what I would have expected from the piece from the comments I've made throughout the assignment ...".*

They used the marking and grading rubric, and once they added up what they had awarded for each section, this then generated the final mark. It was the comments they had written as they had progressed through which cemented the final mark or grade awarded.

'Toni' explores their approach to marking and grading and that of their own department, which explains a community of practice:

*"... I've had to make a conscious change in how a mark because I used to go in with finding all the faults which is what I've just done on the first read through - 'this is wrong, that is wrong and you not done this' but I don't like the nit-picking marking ... because, this was a tip from an external we should be looking for what we can give them credit for it, it's something **[the department]** that we're not renowned for doing, we look for what the student's done wrong rather than what's good ...".*

This excerpt illustrates the different approaches to evaluating an assessment artefact. It seems that lecturers search for things done incorrectly or identification of the things done well, all of which have an impact on the decisions made regarding marking and grading.

'Maz' reports on advice from a much more experienced colleague:

*"... said in her experience there is always a range ... so I actually put them in a kind of a line to see where I am and sometimes I go in and think that is really good but then I find one which is even better and think that's not bad and then you find yourself going back and questioning what you are doing and yeah, yeah it is interesting ...".*

From this advice it appears the sense of having a range or a spread of marks or grades is viewed as just as important as placing a value on each artefact. This approach is related to normative, rather than criterion referencing.



Marking and grading results are shaped into the pattern of a normal data set, known as a 'bell curve', with few marks or grades being seen at the extremes of fail or distinction.

A further influential factor in decisions on the final mark or grade is the moderation process. One interviewee, 'Jae' talks frankly about experiences of their peers and their own attitude toward it:

*"... you know my some of my colleagues, who started at the same time ... about ten of us ... over 50% of us had all come out of practice. And some of them ... would get really get kind of upset ... when they were second marked and the marks were really far apart ... I think marking is very subjective and as long as you justify why you gave that mark then I don't get as upset if that makes sense ..."*

Despite this they remain:

*"... amazed ... from my limited experience how varied it is and how people do it so differently ..."*

From these excerpts it seems that reaching a decision on the final mark or grade to be awarded is fraught with difficulties.

#### 6.3.5.b GETTING A SENSE OF OR OVERALL IMPRESSIONS

This sub theme identifies what getting a sense of the artefact means to those marking and grading. 'Jae' states:

*"... I guess well I feel like I make my mind up on the grade quite early on which I think I get the impression from talking to other colleagues that is quite common ..."*

As they suggest, this is not an unusual position to be in as is intimated by 'Ali' (see 6.3.5.d) and during the protocol analysis interview 'Jae' expressed the following:

*"... I am kind of getting a sense that it's a quite decent assignment already so. I don't want to say I have decided the grade or anything but I'm thinking the first couple of paragraphs are strong so it's giving me confidence that it is a good piece of work ..."*

Without defining the final mark or grade to be awarded, 'Jae' quickly identifies the level of the work they are evaluating. The knowledge of this arises from an analysis of the way the written assessment artefact is constructed:

*"... I think you can ... early on if it is going to be a good assignment generally like you know English and grammar and spelling punctuation is good you know the first reference is correct it is a relevant reference you know within the time frame ..."*

This interviewee goes on to explain other components which present an impression of the assessment artefact including an explanation of how the learning outcomes will be addressed and that it is well worded. Whilst they appreciate this is not a failsafe position, 'Jae' does indicate:

*"... I think generally it has obviously made an influence straight away and I am going to give this grade ..."*

They then go on to relay a story of a very experienced lecturer who would be able to simply review the reading list and from that confirm the grade to be awarded, they were apparently correct in 9:10 cases. Rather than this being the norm, it was a position which Jae denoted as a sign of capability. As highlighted by 'Maz' earlier (see 6.3.2.a) it is often from reading the first one or two pages of an assessment artefact that an impression is gained of its overall quality:

*"... I guess these are the kind of things [**vocabulary and sentence construction**] I am going through in how do you say in the first read that is giving me the first impression ..."*

From these interviewees' extracts it can be seen that how the artefact is presented and structured in the very early stages has a significant influence on the value judgements placed on it by those evaluating it.

#### 6.3.5.c INTERACTING WITH THE ARTEFACT

This sub theme uncovers how the interviewees worked with the assessment artefact to facilitate its evaluation. Illuminating this process are the extracts from four interviewees, illustrating its contribution toward making a

judgement on the value of the artefact. For one interviewee, 'Jae', an initial impression of the artefact was soon dampened:

*"... Clear introduction at the beginning ... the grammar good ... there doesn't appear to be any spelling mistakes. So I am a bit annoyed that they have named the hospital. This is again my experience ... I remember doing a CPD course ... I named the hospital I got an automatic fail for it which I thought was very harsh... But this university isn't kind of like that ..."*

This interviewee is reflecting upon their own very personal experience which is at the forefront of their mind and may have some influence on the mark or grade awarded to the artefact they are evaluating.

'Jo' reviews the assessment artefact to identify if it is as per the task set, which at the beginning, it is clear that guidelines have been followed. What they then come across is:

*"... having to re-read sentences two or three times to get a sense of what it means, which is not a good sign ..."*

Their evaluation continues and identifies:

*"... not using professional language it's not the sort of sentence you would see in an article, some slightly careless errors with their wording though the point being made is good ..."*

This is a second artefact, which at the outset appeared to be of a particular standard, via first impressions. However, for one reason or another continued interaction with the artefact has established a different value judgement.

For 'Ali' their interaction with the assessment artefact leads them to question the content:

*"... which is interesting because I would want him to be questioning why was one antibiotic insufficient, why they give them together ..."*

Questions about the student's knowledge base is soon answered, as they do go on to demonstrate that they understand the why of the proposed treatment, providing some reassurance.

'Maz' highlights how assessment artefacts may be structured in such a way as to not be able to put the most important point first:

*"... [they] will put a really good comment almost at the end of a paragraph or a section ... you can say stick this at the beginning ... that would give you a really good introduction to your paragraph ..."*

This assessment artefact then goes on to show it is:

*"... not quite relating back to the papers [critical analysis of research papers]. So although she has got this theme ... it is not, it's not great. It is not great no ..."*

Whilst 'Toni' relays the following with regard to the artefact they are involved in evaluating:

*"... got quite a narrow scope and considering its postgrad ... expecting them to take a broader view ... now just describing the examination, more describing ... now getting quite bored of reading this introduction or patient history very fine detail ..."*

Interviewee interactions with the artefacts presented here, demonstrate that either expectations or first impressions can sometimes be challenged, with evaluators being left perplexed.

#### 6.3.5.d INTUITIVE DEVICE FOR CONSTRUCTING MARK OR GRADE

All interviewees are required to make use of their institution or department's marking and grading criteria in order to complete an evaluation of the written assessment artefact before them. Interviewees' here talked about how they came to an initial idea about the assessment artefact. 'Ali' talks of their process:

*"... so when I mark I read them through the first time without looking at the marking criteria, I want to get a sense of the overall, my overall impression of it and I think that is informed by the criteria because I know them so well ..."*

In relation to having an impression 'Jackie' discloses the way this is portrayed in their marking and grading process:

*"... I am not sure whether this is right or wrong ... I go through this [the artefact] I give myself all the marks I am thinking and then I add*

*it up cos in my mind I have got a bracket that I want it in. If it doesn't add up to that I would readjust ...".*

These first two interviewees know the existence of a marking and grading criteria but choose in the first instance to get an impression without its use. Once they have gained that impression 'Jackie' proceeds to use the criteria to build a mark or grade that they have decided is appropriate for the assessment artefact.

This position is also portrayed by 'Toni' who concludes:

*"... have got an idea of what sort of marks ... I also think it's funny because everyone talks about the marking grids ... think it's ridiculous to try and pretend that you're not going to put a mark, whereabouts on the grid that it can fall for certain things ...".*

This interviewee believes that, when completing the marking and grading, the grade or level of achievement they have in mind will be engineered by ensuring certain elements on the marking and grading rubric are selected. In relation to the artefact they are evaluating during the course of this interview, the interviewee says:

*"... it seems fairly middle-of-the-road to me so if I get something that's ... not fail but it doesn't feel like it's above 60% so looking to something between 40 and 60 I'm not gonna be looking at the marks I'm just gonna read the comments ...".*

They then proceed to complete the marking and grading rubric:

*"...so can look down the boxes ... so it's sort of split, so I'm going to go in the middle ... which comes out as a total mark of 43 somewhere between 40 and 60 which is where I was expecting it to come ...".*

From 'Toni's' initial intuitive impression they see that realised with using the marking and grading rubric, selecting individual criterion on the rubric which they consider adequately described achievement present in the artefact.

All of these interviewees in one way or another came to an impression of the value of the artefact, which was then born out following completion of the

marking and grading rubric. This could be viewed as a self-fulfilling prophecy.

#### 6.3.5.e SENSE OF DISAPPOINTMENT

This sub theme addresses the feelings of three interviewees, who specifically used this term when evaluating assessment artefacts during their protocol analysis interview. 'Ali' says:

*"... So I have reached the end with a bit of a sense of disappointment. I would have liked him to have done better. He is clearly a very competent clinician in lots of ways but as an academic piece of work it is seriously lacking in skill ..."*

The situation is no better for 'Jacki':

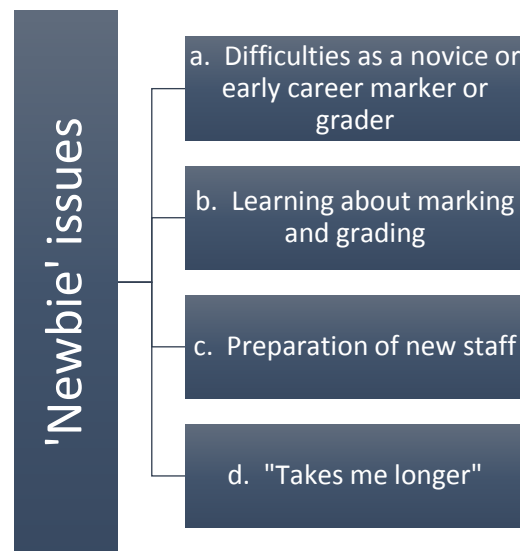
*"... I am slightly disappointed because having read the draft ... my feedback is nearly identical with the draft ... it is a bit disillusioning for me really. Which is why I am huffing and puffing a bit really ..."*

Then finally the extract from the interviewee with 'Jae' states:

*"... I guess now after the good start I have to say I'm disappointed, I've started to feel a little bit apprehensive about it ..."*

All three interviewees are demonstrating their emotional response to the artefact, which they anticipated had the potential to demonstrate good academic achievement. Instead, each interviewee on reviewing their artefact appeared saddened that the threshold set, as in a pass mark or grade, or indeed an expected level of attainment had not been reached. In some cases this had not been without their significant input in reviewing drafts prior to final submission.

### 6.3.6 'NEWBIE' ISSUES



**Figure 17 PA Major Theme**

One objective of this research was to explore novice lecturers' thoughts on marking and grading to increase understanding of their participation in evaluating written assessment artefacts. To fulfil this aim, lecturers with 2 years or less experience in higher education were deemed to be novices and as such were specifically targeted for recruitment at all participating institutions.

This major theme specifically addresses concerns raised by those relatively new to marking and grading, referred to in this study as 'novice' markers or graders. Data presented below is from three early career academics and demonstrates their anxieties and sources of support.

#### 6.3.6.a DIFFICULTIES AS AN EARLY CAREER MARKER OR GRADER

This sub theme begins to explore the lived experiences of those thought to be novice markers or graders. Excerpts here identify some of the difficulties associated with the position of being a 'novice'. 'Ali', who had been an academic for two years, speaks candidly of their experience:

*"... I think when I started I found I really struggled with it, ... I was kind of looking for this magic formula that somebody somewhere would*

*know I kept asking everybody "How do you mark?", well you just follow this. Well what is understanding, what is analysis ...".*

The questions posed by 'Ali' were felt to be those to support their understanding of what it was exactly that they were supposed to be evaluating in the written assessment artefact. One of the potential issues was what it looked like when these elements were present, 'Ali' then goes on to disclose how they came to being able to undertake the required evaluation with more confidence:

*"... I think what really helps is uhm, we have just restarted like the old marking workshops where we will all just mark the same piece and then come together and talk about it. And that is brilliant; if they had had that when I started I think it would really really have helped ..."*

For 'Maz', 18 months into their academic career, they consider:

*"... nothing really nothing prepared you for that first time when you mark. One of the first things I marked here was we had a seen exam ... and so I can remember sitting at home with them going through them and through them and thinking I don't know what I am doing here ..."*

This interviewee is expressing feelings of inadequacy in relation to being able to fulfil their role of evaluating the written assessment artefact, a seen exam paper. They go on to qualify that statement:

*"... everybody else seems to know what they are doing, well I mean the rest of the team have been here around for a while so I went off to talk to others in different areas of the faculty who were new or we had what we called a 'newbie group' ...".*

'Maz' does accept that though they still feel to be a 'novice':

*"... you probably build up your expertise during the marking period and then you kind of relax and the next one comes along and you actually ooh ooh can I remember everything ... give[ing assessment students] them lots of feedback and lots of support in between so I think that's growing my feedback skills ...".*



A number of other new staff had started at around the same time, and they formed a group to provide support to one another. Whereas for 'Ali' support in their evaluative role for assessment artefacts could have been improved much earlier by participation in team marking and grading workshops:

*"... you are learning from wisdom from other people ... we did one recently I was sort of 20 percent lower than somebody else. So initially I was like Oh god I am really stingy but then as we talked about it and teased it out they were like Oh actually there is room to move in both ends ..."*

The impact of such interactions is a significant increase in confidence. Whilst the story from 'Jae' was very different:

*"... well I probably shouldn't say, but I think I have not had a great year and a half here ..."*

These interviewees cite support from others including more experienced academics in marking and grading as an important component of increasing knowledge, skills and confidence in their evaluative role.

#### 6.3.6.b LEARNING ABOUT MARKING AND GRADING

The interviewees spoke of specific support or development they had participated in which assisted them in learning the craft of marking and grading. This excerpt from 'Maz' identifies advice they were given on the Post Graduate Certificate in Education:

*"... When I did my PGCE the general advice was you didn't put somebody too close to the top or the bottom of a range ..."*

This has caused this interviewee some difficulty, feeling that not all marks or grades could be in the middle of the range yet with this advice:

*"... to make sure it is either a clear pass or a clear fail but **[identifying]** where it sits within the criteria that is what I still find quite difficult ..."*

From this extract it appears the educational programme, rather than improving performance of their marking and grading role, has increased uncertainty with regard to the parameters of a 'clear pass or clear fail'.

This seems to be a similar story to that of 'Jae' who feels there was little information:

*"... the academic development programme, but as a department we have not had any workshops or anything like that it was quite have a go see what you think ..."*

Their final summation of their current situation is that:

*"... at the moment I am not, I don't really feel very well supported you know ..."*

It is difficult to hear that an academic reaching the end of their first full academic year in higher education should consider that support and preparation for one of their most major roles is deficient.

The final 'novice' interviewee, 'Ali', explores their preparation for marking and grading slightly differently, they analyse the task which they are asked to do by questioning whether it is possible to evaluate an artefact:

*"... trying to capture practice, in something that is reducible to a tool and I don't think with marking that you necessarily can ..."*

For 'Ali' the disorderliness in the context of clinical healthcare practice raises some difficulties in trying to capture student achievement within the confines of a marking and grading rubric. Despite their 'newness' they consider:

*"... I do think you can recognise quality. And I think a lot of it is about trusting, trusting what your impression is of it. So the intuition is there from the word go ... it is different in healthcare in respect of you coming into it with already a sense of quality ..."*

For 'Maz' the fact that they have used their evaluative skills in their previous role in clinical practice, is seen as a transferrable skill. Despite this seeming level of confidence, this interviewee then goes on to say (at the completion of the protocol analysis interview):

*"... I don't think you will learn anything from me I am too much of a newbie ..."*

That feeling of being newer than other colleagues, thereby having nothing of interest to contribute, pervades.

### 6.3.6.c PREPARATION OF NEW STAFF

One interviewee, 'Jae' highlighted their poor preparation to become effective in marking and grading and they ensured that heads of department knew how they felt. However for 'Maz', there was a different experience and this was through the availability of a mentor. This was a very important contribution to their development:

*"... I did get one or two marked and she had a look and second marked them and then we compared comments and so I have been very well supported in ... really I have to say ..."*

Another interviewee reflected on their potential approach to support a new member of staff:

*"... If I was going to talk to someone who was new about it I would probably give them examples like I do with my students Give them an example of 10 percent essay and a 40 percent essay and a 60 and an 80 that have been checked that have been second considered and moderated so it is a kind of fairly reliable indicator ..."*

This interviewee, 'Ali', had not experienced anything like this until department marking workshops took place, after they had been given a number of assessment artefacts to evaluate.

'Maz' does highlight an interesting aspect, which is as a new lecturer, where moderation takes place they take a back seat or are even reticent in voicing opinions:

*"... I think when you start you think nobody's going to agree with me but actually you find you are not far ..."*

A further approach which is adopted to introduce new lecturers to different marking and grading styles or approaches, is to embrace a different constitution of teams who will complete first marking and second marking. 'Novice' markers are thus exposed to different marking and grading conversations, especially during moderation processes, serving to develop their practice.

#### 6.3.6.d “TAKES ME LONGER”

The final aspect commented on by those new to marking and grading was related to the time aspect. ‘Jae’ comments:

*“... when I have got a new set of marking I think it takes me longer but then I’ll start copying and pasting some phrases from some other ones that relate to the assignment ...”*

This step facilitates ‘Jae’ in constructing feedback that is comparable across criteria on the marking and grading rubric. They do concede:

*“... as I have only just started I know that it will take me longer than perhaps if I am in the swing of things ... I know I take a lot more a long time marking. It is something I need to get better at ...”.*

As a final observation from ‘Jae’, they mention that their institution want staff to improve the marking and grading in response to internal and external drivers (feedback from surveys), their observation is:

*“... we are only allowed to spend this amount of time on it so you know you can’t have it both ways can you ...”.*

For ‘Jae’ a reduction in the amount of time available to complete the evaluation of written assessment artefacts will not result in an improvement in marking and grading.

‘Maz’ uses their time in a different way when evaluating an artefact, they read it through once, identifying general issues and making general comments:

*“... go back to try and make sure I have got a real take on the actual structure a little bit more and that may be dreadful use of time but it is the only way I can do it, I don’t think I can give any 4,000 word essay credit by just having a quick whizz through ...”*

Time for ‘Maz’ may be an issue, but they don’t feel that they can undertake a careful evaluation including constructing useful feedback any quicker.

This major theme has explored the experiences of three ‘novice’ lecturers, from three different institutions, in developing their skills and competence in the evaluation of written assessment artefacts. The themes facilitated

identification of factors that aided lecturers along with those which, from interviewees' perspectives, hindered their development in this aspect of their role.

### 6.3.7 SUMMARY

Completion of interviews using the Protocol Analysis method has facilitated intimate inspection of what actually occurs during the marking and grading processes undertaken during a lecturer's interaction with an assessment artefact in HE. To prepare the data for analysis the steps identified in Chapter 2 for application of framework analysis were instigated. Data presented represents that which could be used from eight of the 11 interviewees. Reasons for this are discussed in full in Chapter 8, Limitations (see 8.2). Extraction of useable data provided opportunity for the construction of six major themes and embedded sub themes giving a rich description of the lived experience of lecturers, in action, during the marking and grading process. Findings retrieved via this interview method are reviewed in light of contemporary knowledge, contextualising the subject of marking and grading and are presented in relation to the research objectives and research questions (Chapter 7).

A post protocol analysis interview was completed for each interviewee on completion of their interaction with the assessment artefact which provided opportunity for both parties to seek clarification on any aspects. This also provided a debrief opportunity for interviewees, it was seen an important step ensuring each individual felt comfortable with the interview process given the usual solitary nature of marking and grading activity. This aspect is explored further in Chapter 8, Reflexivity (see 8.3), with cognisance being taken of my role in the interview process and then following this in the construction of the story of marking and grading being told here.

# Chapter 7 FUSION OF FINDINGS

*“None of us are to be found in sets of tasks or lists of attributes; we can be known only in the unfolding of our unique stories within the context of everyday events”.*

Vivian Gussin Paley, (1990)

The penultimate chapter will return to the overall outcomes from the data exploring how well they have provided answers to the research questions. It is also important to examine how understanding of the data has assisted in moving towards achievement of the research aim. This aim was to investigate the process of marking and grading from the perspective of the marker. To recognise good marking and grading practice, deriving a model of ‘expert’ practice aiding the development and proficiency of novice markers.

This chapter is the first time that findings from the two data collection methods have been brought together to develop integration and coherence from the major themes of the earlier chapters. This brings together the findings from the PA and CI methods used to study lecturers’ approaches to marking or grading of written assessment artefacts. Combining the findings from these methods is paramount and has not been attended to earlier in this thesis. A systematic approach for the exploration and amalgamation of findings has been achieved by utilising the tools which framed the findings, namely the research objectives and research questions. This is the first step towards beginning to demystify marking or grading processes and exploring any evidence for the development of expert practice. The chapter develops by relating findings to the contemporary literature which provides supporting or opposing evidence. Further organisation of the findings is achieved by using the four study objectives along with the research questions presenting an analysis of evidence from the interviewees via the two research methods. A table, identifying where evidence for each objective was obtained, is presented at the start of each separate section to facilitate a return to full excerpts from interviewees.

The pragmatic decision to utilise framework analysis to extract findings from both interview methods has facilitated the integration and presentation of this evidence here. Each of the four research objectives are interrogated by application of the three research questions (Table 5). Lastly significant outcomes are identified at the end of each research objective, these are then utilised to support final conclusions as research outcomes, prior to further exploration and suggestion for application in the final chapter.

<p><b>1. Examine assessment by exploring marking or grading practices.</b></p>	<p><b>2. Identify what are 'good' marking practices to develop a concept of 'expert' marking and grading</b></p>	<p><b>3. Explore 'novice' lecturers thoughts on marking or grading</b></p>	<p><b>4. Explore cognitive processes and extraneous influences on marking or grading practice.</b></p>
<p><b>What do lecturers believe they do when they are undertaking marking or grading?</b></p>			
<p><b>What thought or cognitive processes occur in lecturers during the act of marking or grading?</b></p>			
<p><b>What exists to ensure equity or consistency in professional or inter-professional marking or grading, this includes inter-rater reliability.</b></p>			

**Table 5 Research Objectives and Questions**



## 7.1 RESEARCH SYNTHESIS

This study and analysis is based on exploring lecturers' processes for evaluating a written assessment artefact. This synthesis includes detail of lecturers' approaches to making a final judgement as to the value of the written assessment artefact in terms of allocating a mark or grade. The three research questions are central to identifying how lecturers approach, create, and produce the final assessed artefact for return to the student, the development of which is a puzzle to those new to the profession. Therefore processes for supporting the development of lecturers into someone, who can mark or grade a written assessment artefact appropriately (as is defined by the profession at the current time), thus appears mysterious. As is hinted at by Paley's (1990) quote above, close examination and exploration of lecturers' approaches to marking or grading are the only way to discover how lecturers become markers or graders and how that becoming is framed. Exploration continues toward identifying how this new knowledge could then be shaped to support those new to the profession in a more constructive and purposeful way.

The two research methods used have revealed different features all of which are necessary components of the process of marking and grading. Protocol Analysis (PA) was effective at revealing exactly what lecturers did during the whole process of marking and grading. On the other hand Cognitive Interviewing (CI) to a greater extent revealed lecturers' espoused approach to the practice of marking and grading. The remainder of this chapter is focused on exploring the findings for each of the four research objectives, framed in turn by the three research questions whilst being illuminated by contemporary literature, where this exists. Major themes have been used to present the research synthesis, with a table being produced for each research objective illuminating which sub theme the interviewee quotations originated from. This approach was felt to further underpin my approach to rigour relating to dependability, credibility, transferability and confirmability of my findings.

### 7.1.1 OBJECTIVE 1: EXAMINE ASSESSMENT BY EXPLORING MARKING AND GRADING PRACTICES

<b>QUESTION 1: WHAT DO LECTURERS BELIEVE THEY DO WHEN MARKING OR GRADING? (CI)</b>	<b>QUESTION 2: WHAT THOUGHTS OR COGNITIVE PROCESSES OCCUR IN THE LECTURER DURING THE ACT OF MARKING OR GRADING? (PA)</b>	<b>QUESTION 3: WHAT EXISTS TO ENSURE EQUITY OR CONSISTENCY IN PROFESSIONAL OR INTER PROFESSIONAL MARKING OR GRADING, THIS INCLUDES INTER-RATER RELIABILITY?</b>
<b>Settling Down to Do The Job</b> 4.3.1a, 4.3.1b, 4.3.1c, 4.3.1d	<b>Student Assessment Literacy</b> 6.3.2a, 6.3.2b, 6.3.2c	<b>Settling Down to Do The Job</b> 4.3.1a, 4.3.1b, 4.3.1c, 4.3.1d
<b>Parameters of Practice</b> 4.3.2c, 4.3.2d	<b>Operational Necessities</b> 6.3.3a, 6.3.3c, 6.3.3d, 6.3.3f	<b>Operational Necessities</b> 6.3.3e
<b>Messiness of Marking or Grading</b> 4.3.3a, 4.3.3b, 4.3.3c	<b>Tangible Marker or Grader Actions</b> 6.3.4a, 6.3.4b, 6.3.4c	<b>Parameters of Practice</b> 4.3.2c, 4.3.2d
<b>Knowing</b> 4.3.4a, 4.3.4b	<b>Implicit Knowledge Or Actions</b> 6.3.5a	<b>Messiness of Marking or Grading</b> 4.3.3a, 4.3.3b, 4.3.3c

Table 6 Objective 1

## **Question 1 - What do lecturers believe they do when marking or grading? (From CI interviews)**

### 7.1.1.a SETTLING DOWN TO DO THE JOB

One of the opening questions to interviewees participating in Cognitive Interviews (CI) was getting them to reflect on the first things that go through their minds when they come into contact with an assessment artefact. One interviewee 'Eddi' (CI) considered spelling and grammar errors to give an immediate poor impression.

This initial impression is also expressed by 'Harri' (CI) and other interviewees and although they try and remain bias free, acknowledging that is not always possible.

The question of what interviewees did and when was subject to investigation, as it was with Tomas (2013) study. For instance 'Gerri' (CI) identifies their starting point, which is administrative to make the marking or grading reporting an easier process ensuring that they set up files on their computer incorporating the appropriate student details.

Many interviewees made notes as they were working through the assessment artefact. This was a purposeful act as 'Alex' (CI) indicated it prevented them from getting some way into the evaluation and then being unable to remember what they had read early on. To combat this they made notes consistently throughout considering the whole artefact. Interviewees were concerned with is ensuring that they know what the students were asked to do.

Interviewees talked of how they work with a batch of assessment artefacts, they do not make reference to the utilising of marking or grading criteria. This approach fits that which was found by Bloxham et al. (2011), when their participants used the marking or grading criteria post evaluation to justify the decisions they had reached. Here, 'Gerri' (CI) is an experienced evaluator of assessment artefacts felt confident to note the mark or grade as they proceeded demonstrating less reliance upon the written criteria present on a marking or grading rubric to make their initial judgement, a situation which is not unknown within the literature (Brooks 2012).

As interviews progressed it became clear that there were different processes in action in relation to how marks or grades were attached to an assessment artefact, 'Alex' (CI) proposed that they review the artefact for where they could award marks.

The above extracts show how evaluators allocate marks or grades is considered by identifying two approaches. 'Alex's' (CI) comment relates to how the artefact itself is structured thus enabling effortless identification of the assessment task requirements. Debates about students' abilities to communicate through the medium of writing have continued to be explored (Lea and Street 1998; Lillis and Turner 2001). Hunter and Docherty (2009) explore differences in evaluation of student writing from the perspective of the lecturer by identifying the tacit beliefs of those marking or grading assessment artefacts.

The significance of student writing and its presentation within the artefact and therefore by implication evaluator judgement of achievement is perhaps not stressed enough to students. Research undertaken demonstrates that presentation does play a factor in the mark or grade achieved by an assessment artefact (Hartley et al. 2006). Focus on developing writing for the academy, including the particular approach required by the academic discipline may facilitate improvements in artefacts submitted for assessment. Contemporary evidence does not present sufficient indication of this being the entire explanation of inconsistencies in assessment artefact evaluation (McConlogue 2011; Bloxham et al. 2015).

An additional consideration voiced by interviewees was in relation to their existing knowledge of the topic, seeing this as an advantage 'Kim' (CI) highlighted that it was an advantage when the artefact was within their practice area as they possessed appropriate knowledge.

Pre-existing knowledge of the artefact made its evaluation a much easier process. This interviewee as do others comments about the ease with which the artefact could be read, due to it being well written, good spelling, grammar and punctuation including being succinct with an effective structure. This latter observation could have impacted upon the ease with which its

evaluation was completed, rather than solely previous knowledge or expertise in the topic area. When this was not the case the poor presentation made evaluation of assessment artefacts become a much more difficult process.

'Gerri' implements a particular process to enable them to be satisfied that they have allocated appropriate marks or grades via a second level of interaction with the artefact. They discussed reviewing the artefact again to ensure they were confident that the script with the highest mark really was the best.

Reviewing it with the others ensuring that it had been allocated what was warranted amid the other artefacts in the batch being evaluated. The practise was essentially completing a process of normative referencing in addition to criterion referencing which Lok et al. (2016) urge should be viewed as less in opposition and on more a continuum, with one informing the other.

Interviewees suggest a review or read of the marking or grading criteria establishes some parameters to assist them in identifying the components of a 'good essay' and establishing a baseline for a failing essay. Interviewees hinted that this is an activity they undertake whilst preparing to evaluate the artefact, fitting with recent findings (Tomas 2013). Whilst other interviewees were aware of the presence of marking or grading criteria and knew it would assist in maintaining an objective rather than taking a subjective stance, 'Gerri' (CI) reported the use of the submitted essay, comparing one with another. Within a group of artefacts they needed to be assured of how the artefacts related to one another, evidencing interviewees employing normative rather than criterion referencing (Lok et al. 2016).

Occasionally evaluators will read all or part of an artefact when completing the marking or grading rubric 'Chris' CI hints that they review the grid and parts of the essay. This double checking in this instance is a method of increasing surety in their evaluation numerically (or alpha) and qualitatively through comments.

'Gerri' (CI) highlights a difficulty with distinguishing clear differences in essays only a few percentage points apart which concurs with Sadler's (2009b) opinion on the difficulties of using any type of codification to describe achievement. Essentially this interviewee is highlighting the inherent difficulties with being able to adequately explain, qualitatively, what the differences of achievement would be within what appears to be a fine gradient in the two assessment artefacts.

'Terri' (CI), reflecting on their marking or grading practice, identifies that knowing what to do is as important as the know-how, bringing into focus the potential utility of being part of a community of practice (Wenger 2011). Being party to the conventions within the department for use of the marking or grading tools available would enable the evaluator to ensure they did not transgress any unspoken rules of engagement for example annotation of a script. Issues such as these are important considerations to ensure effective team work, but increase evaluator confidence in the task before them, whether they are new to the role or new to the department.

Whilst a different perspective of the utility or impact of the rubric was raised by 'Gerri' (CI) who felt that using criteria pushed evaluators into a negative mindset, identifying gaps in the student submission.

This raises the issue of the criteria making evaluators search for negative aspects within the artefacts which would lead to a different outcome compared to a tool guiding evaluators to search an extract for positive attributes.

#### 7.1.1.b PARAMETERS OF PRACTICE

This research identified a practice that interviewees use to support and induct 'novice' evaluators and that is double marking or grading. It is a process which is known to be time consuming, but this was viewed as an alternative to trying to sit with a new evaluator going through a step by step walk through of what to do.

The adoption of such processes will not be overtly considered when departments are attempting to aggregate workloads amongst its staff in

times of resource restriction (Burgess 1996; Burgess et al. 2003).

Contemporary approaches to resource reduction is to utilise what is currently available more efficiently. Lecturing staff account for a large proportion of institutional resource, there is a need to manage this more effectively, hence introduction of workload models (Burgess et al. 2003). Such models are used to identify individual staff workload for every aspect of their three roles, commonly divided into teaching, administration and research. Proportional times are calculated for every activity, including marking and grading of assessment artefacts. This will facilitate a time allocation for completing the marking and grading process, from first contact with the artefact through to completion of feedback for the student. This time allocation will be outside consideration of the status of the academic (experienced or novice), or the support being provided to others during the process.

However, for 'Harri' (CI) it was their revelation that they possessed limited confidence in their own ability to mark or grade that then made it difficult to articulate their processes to others. An educator knows that a student understands a topic when they are able to describe it in their own words. It appears for 'Harri' that this is not the case for marking or grading.

Interestingly this interviewee reflects on how they learned how to undertake marking or grading, this consisted of another evaluator talking to them. They do not define what this consisted of, or if it was similar to the process 'Harri' had adopted for the preparation of new staff:

*“... you latched on ... learnt their ways ... whether their way was right or not who knows ...”*

This excerpt identifies that they are still concerned with the correctness of the process (Annala and Mäkinen 2016). This and other excerpts demonstrate that interviewees across a variety of institutions feel under prepared for their evaluator role in which they experience a dichotomy between currently being a 'novice' academic and expert practitioner in a previous role (McDermid et al. 2016). 'Harri' is an academic with many years in the academy, but had to have had previous years in clinical practice before their conversion to academic. Many health academics draw upon their previous clinical practice experience for developing and delivering teaching and learning to students,

but this knowledge is often inadequate to support their new evaluative role in HE (Boyd 2010).

When exploring how an evaluator completing marking or grading knows that they are right, comes down to experience. 'Billie' (CI) considers how the kind of experience is important and the incorporation of that by the individual becomes a tacit action, potentially relegating it to being unexplainable (Duguid 2005; Entwistle 2008). This is not to say that this knowledge is not significant but that individuals find articulating what have turned into implicit and seemingly automatic actions difficult to explain. There has been a suggestion that the apprenticeship type model of learning has thus been successful as knowledge about the area is transmitted in two ways, through explanation and observation. The inner world of an evaluator marking or grading is very rarely observed by others, especially inexperienced or 'novices' or those new to a marking or grading role.

Where such observation would be useful would be for instance in how often evaluators review with a single assessment artefact. Research by Bloxham (2011) and Tomas (2013) both highlight that it is not uncommon for evaluators to require more than one contact with an assessment artefact in order to be able to make a judgement as to its value. Conversely 'Chris' (CI) states:

*“... I rarely read it again ...”*

'Chris' is an experienced evaluator with over 20 years' experience in higher education who demonstrates that experience plays an important part in the development of confidence in executing their role.

#### 7.1.1.c MESSINESS OF MARKING OR GRADING

Through reviewing interview extracts it appears that a number of experienced interviewees have a sense of what the artefact will be awarded in terms of a mark or grade prior to completing the evaluation process. This has been termed intuition or insight in this research, and it occurs in the initial stages of interacting with the artefact. Initial impressions serve to act as a barometer from which to seek confirmatory evidence within the artefact, as



'Alex' (CI) identified they were confident when the mark or grade achieved through application of the rubric ended up in the same '*ball park*'.

This fits with the approach proffered by Hand and Clewes (2000) who identified that evaluators did not consistently utilise criteria or guidelines which were known by other staff and students. At these times lecturers appeared to be reliant upon tacit knowledge of the requirements for the assessment artefact.

Some interviewees expressed such confidence in their initial insights that they were prepared to make changes to the rubric so that these matched. Further confirmation for this approach is provided by other interviewees e.g. 'Chris' (CI):

Yorke (2011) highlights the presence of objectivity in the act of marking and grading assessment artefacts as a misleading notion. They identify that evaluators commonly use two approaches, namely holistic and criterion referencing, to achieve a judgement regarding, an assessment artefact. The existence of being uncertain regarding judgements was highlighted by Baume et al. (2004) when lecturers seemed to be wavering amongst their overall holistic judgement and what the criteria for marking or grading are indicating. Once a group of artefacts have been evaluated, a process of norm referencing was also described by interviewees. These approaches have also been validated by other authors (Sadler 2005; Carless 2006; Hodgkinson et al. 2008; Sadler 2009a; 2009b; 2013).

Evidence suggests that the current assumptions as to both actions and motivations of evaluators are poorly understood. Along with this Brooks (2012) gathers evidence from a number of studies asserting that where uncertainty exists lecturers are susceptible to peripheral factors which can lead to an increased level of subjectivity in the evaluation. The approach of the two interviewees above raise further questions in relation to the place of intuition or tacit understanding in the allocation of marks or grades for an assessment artefact. . The ability to use insight or intuition is not available to all evaluators as 'Harri' (CI), in common with others, considers that this is developed through years of practice. This position was supported by other

interviewees in my research cohort. They expressed the idea that being able to complete a holistic assessment of an artefact was something that developed with more frequent or prolonged exposure to evaluative opportunities.

The argument for removing the need for a fine grade description of achievement when evaluating an assessment artefact presents itself when considering 'Terri's' (CI) perspective. They believe that evaluating an essay bears no resemblance to assessing an artefact for which there is a correct answer and therefore has an accurate mark or grade which can be awarded. They suggest at best deciding on a grade band, in which the value of the artefact lies, should be sufficient. This makes the case for the use of more holistic evaluations of assessment artefacts as suggested by current literature (Yorke 2010; 2011).

A number of interviewees discussed workload in relation to increasing fatigue, shorter deadlines and delivering on academic responsibilities all impacting on the task of marking or grading. Saunders and Davies and Davies (1998) identify the possibility of marking or grading being completed with less rigid adherence to the assessment criteria, increasing inconsistency in artefact evaluation. Two interviewees in particular ('Billie' (CI) and 'Gerri' (CI)) considered the effect of evaluator fatigue and the need to pace the work. Others ('Lesli' (CI) and 'Eddi' (CI)) identified that this may mean eroding own time to enable deadlines to be met. This effectively portrays interviewees' perspectives on the imbalance between workload and work time available to complete the job requirements resulting in interviewees needing to use what would be leisure time to complete the task.

These aspects of workload do not appear to be taken into consideration when institutions agree to speed up the processing of assessment artefacts, at the request of students (en mass). 'Danni' (CI) expressed feelings of a number of academics, in that marking and grading could not be completed any quicker. They also wondered what the outcome would be if they could canvass student opinion on which they would find more reassuring, speed or the completion of a rigorous marking and grading process.

Lastly 'Harri' (CI) summarises this argument rather eloquently:

*“... there's so much rides on it ... I do take ... more time that I should ... it's too important to students ...”.*

The importance of assessment is recognised by the Quality Assurance Agency for Higher Education (QAA 2012). This document sets out the importance of assessment in establishing and maintaining standards of awards for programmes of education in HE institutions.

A further concern, voiced by interviewees, was being able to select with accuracy the mark or grade which reflects the value of the assessment artefact. Sadler (1987) first mooted the difficulties of applying criteria in order to establish the value of an artefact nearly 30 years ago. They define what they see as difficulties with applying a fine grade verbal description which are turned into individual criterion requiring selection by the evaluator in order to describe the achievement. An example given by Sadler is in relation to coherence and being able to explain the difference between reasonably coherent and highly coherent. This is based on an assumption that others have the same understanding and will thus employ this interpretation in the evaluation of the artefact. Twenty years on Sadler (2007) continues to expose what they see as the continued dangers presented with breaking both learning and assessment down into ever smaller, more easily digestible chunks. They view this as diluting the overall endeavour, to increase capacity for learning. It seems that 'Gerri's' (CI) assertion above, hints at the difficulty in applying imperceptibly granular descriptions of differences in learning, without that component itself being reduced down to small countable but potentially meaningless fragments, is true. This position calls into question not only the accuracy but also the validity and reliability of evaluation of the assessment artefact.

A number of interviewees in different ways explored processes of norm referencing, where the achievement of a particular assessment artefact is reviewed in light of others in the cohort. This is a way of ensuring that there is a spread of marks or grades across the cohort, with limited skewing either up or downwards. Such a measure can be obtained by reviewing both the mean mark and standard deviation after the marking or grading is complete.

However, interviewees appeared to want to ensure that their group of marks or grades were not viewed as being overly stingy or generous in comparison with their peers.

They use the evaluation of this assessment artefact as the barometer from which to judge achievements. The effect here is to try to objectify the marking or grading outcomes by presenting the assessment data as a normally distributed data set. In truth it would be unusual to find a group of assessment evaluations which followed this convention. Where this would be possible is when each grade band or mark has an allocated quota to be awarded (Leathwood 2005). Such an approach does not present effective evidence of the reliability or validity of evaluation procedures, with artefact performance being judged in and amongst the cohort with which it was submitted. This is of benefit in a year when achievement has not been of a particularly high standard, but is a disadvantage within a stronger performing group or cohort.

#### 7.1.1.d KNOWING

Interviewees reflected on assessment artefacts which had been successful and unsuccessful. They postulated reasons for non-achievement of the requirements and it was often related to poor author assessment literacy. Disappointingly this often seemed to be carried forward to supplementary assessment stages despite advice, for example 'Chris' (CI) recognised that much advice given to one student had not been actioned.

The impact of efforts to secure development of appropriate skills in assessment by publication of evaluative criteria and standards along with assessment task requirements appears to be limited. Within the literature there is a wealth of evidence validating this current position (Taras 2002; O'Donovan et al. 2004; Sadler 2009c; QAA 2012; 2013; 2014). Interviewees appeared to do all they could to try to identify if the assessment requirements had been achieved including feeling that they may need to read the work with more astuteness to identify success.

Interviewees here expressed their desire to review these artefacts as positively as possible, willing evidence of achievement to appear, (Lea and

Street 1998; Grainger et al. 2008), as evidence assures evaluators of the utility of previous assessment feedback (Higgins et al. 2002).

Interviewees suggest an important aspect preceding any evaluation process is that of knowledge of the assessment task itself. It is suggested that for evaluators to complete their task effectively they needed to have an implicit sense of what the artefact would look like, and know what they would be expecting. In the realms of knowing, tacit knowledge comes to the fore, incorporating knowing that with knowing how (Wyatt-Smith et al. 2010).

'Billie' (CI) refers to this as:

*"... propositional knowledge ... you know how, but ... can't always explain how ..."*

This becomes an issue when trying to transfer knowledge across marking or grading teams to develop a community of practice (Herbert et al. 2014) in an effort to share tacit interpretations.

Interviewees in one way or another when discussing how they came to a judgement describe a holistic evaluation of the assessment artefact (Sadler 2009c; Sadler 2009d). Various indicators are highlighted as important in this process including the way the essay was written ('Danni' (CI)). As Brooks (2012) identifies it is not only the criteria that influence the mark or grade but the existence of a stereotypical assumption of an individual or groups abilities. At the outset such indicators may signpost the final judgement as evaluators begin to consider the worth of the artefact. Their thoughts are supported by an objective application of the marking or grading criteria.

Evaluators decisions are based on applying objectivity gained by using a marking or grading rubric, supporting findings from Bloxham et al. (2011), which suggest that evaluation of artefacts are often completed in a two stage process. The initial stage coming from a holistic evaluation referred to by Bloxham et al (2011) as an *'initial judgement'* (p 662), followed by a second stage in which when evaluators refer to the assessment criteria or grade descriptors to facilitate them honing their decisions. What was noted in their study was that a clear differentiation between stages could not always be

noticed, with a move straight to allocating marking or grading decisions very quickly from the 'initial judgement' stage.

## **Question 2 - What thoughts or cognitive processes occur in the lecturer during the act of marking or grading? (From PA interviews)**

### 7.1.1.e STUDENT ASSESSMENT LITERACY

Interviewees' were concerned with academic style and revealed the importance of this in presenting a good impression of the assessment artefact. Interviewees presented the case for a well organised and well-articulated artefact and how this heavily influenced the perception of its value. As 'Maz' (PA) confirms earlier findings when an essay is well structured it is much easier to read and therefore to allocate a mark or grade.

Organisation appears to be key, and it should potentially be highlighted to students in preparation for assessment. As another interviewee suggests, better organisation would lead to less frustration in terms of the individual completing an evaluation (see 6.3.2a - 'Shirley' (PA)), by having to continually consult the appendices.

Though interestingly for one interviewee, the impact of the first assessment artefact which they come into contact with, in some way seems to set the scene for subsequent evaluations. This for 'Toni' (PA) leads to a number of suppositions not really knowing how 'bad' a 'bad one' was until in their own mind they have sorted out "... *a marker post* ...". This being one where they can establish a benchmark for what they have viewed as good or satisfactory, as excellent or as a failure to comply with task requirements. This judgement therefore is made in light of knowledge of other assessment artefacts being evaluated, providing further evidence of the role of norm referencing, and is in keeping with contemporary research (Bloxham et al. 2011). The concept of heuristics is identified (Brooks 2012) as a method utilised by evaluators to compare recently evaluated artefacts in current assessment batch or to the students own previous answer (e.g. short answer assessment paper). These approaches facilitate evaluators reaching the judgement required to award a mark or grade. A number of authors identify

that the use of norm referencing is more common than is potentially publicly known in an era where criterion referencing is held in higher esteem (Saunders and Davis 1998; Leathwood 2005; Sadler 2009b). This demonstrates the difficulty with adopting Husserl's position of '*Epoche*', detachment from what is known, in order to make an evaluation of the assessment artefact.

When evaluating assessment artefacts, interviewees talked about wanting to be able to identify analysis. The presence of analysis is formed through an impression of the arguments that are presented and the standing of the evidence which is used to support them (Thompson 2001). The artefacts which lacked an analytical approach could be identified very early into review the artefact, which could be on page 1. This for 'Ali' (PA) evident through that the artefact being lacking in evidence of presenting a position and then identifying how it was intended to be explored, thereby resulting in unmet expectations.

Poor writing technique then leads to a number of problems with evaluating the artefact as already alluded to. It appears that the impact of writing style and academic proficiency in the use of language cannot be understated. Whilst the over use of quotations was a concern for interviewees, they questioned if students knew the purpose of a quotation, again 'Ali' (PA) considers that they should save them for when:

*"... there's no better way of saying it ..."*

One interviewee also highlighted the issue of what mark or grade you award and the message that this sends to assessment students who may misunderstand the significance. A mark or grade which uses percentages (a fail 39% or less) or grade descriptors (e.g. 'marginal fail') could still convey the impression that there is little to do to correct it (see 6.3.2.c).

#### 7.1.1.f OPERATIONAL NECESSITIES

An interesting aspect explored was how interviewees dealt with unknown or unfamiliar assessment topics and how they found it difficult to evaluate when their underpinning knowledge was limited compared to a subject they knew

well, 'Jae' (PA) identifies that they feel more confidence and less apprehension when marking and grading a topic they know.

Anxiety continued when interviewees were unaware of the structure and purpose of the assessment artefact. Knowledge of the task was paramount to act as a benchmark from which to undertake the evaluation as is highlighted by Sellbjer (2015), who contends that different interpretations of assessment tasks presents a challenge. Those evaluating assessment artefacts will unintentionally provide a different mark or grade and feedback all connected to their level of understanding of the assessment task.

Judgement of the marker or grader comes into play when there are no explicit guidelines regarding weighting of elements within the artefact, meaning they must decide what is the most important. Knowing how their particular departmental colleagues would approach this allocation is imperative to maintain a unified approach to the evaluation. Along with this comes the fine-grade decision of how many marks or what grade to award for certain elements or criterion, for aspects of the artefact or as a whole. In essence this is related to use of criteria to generate the mark or grade

Whilst 'Shirley' demonstrates a leaning toward normative based referencing, but in part due to having had contact with a batch of artefacts and by the final artefact they started comparing one with another. For this interviewee this appears to resemble a final checking process increasing their surety that they have awarded the right marks or grades to the artefacts they have been in contact with. This in some ways resembles the approach of 'Maz' (PA) who wants to put all the artefacts in a line worst to best.

During the PA interviews it became evident that there were times when interviewees had some difficulty on deciding upon a precise level of achievement of assessment artefacts. This was most apparent when they were considering an artefact which was either on the cusp of achieving a higher mark or grade ('Jae' (PA) 6.3.3.c) or indeed one with the potential of being awarded a fail ('Ali' (PA) 6.3.3.c). It seems there are times when the level of achievement is very difficult to quantify. Some authors would argue



against the use of such narrow restrictions which are present when using percentage to quantify achievement (Yorke et al. 2002; Yorke 2010).

#### 7.1.1.g TANGIBLE MARKER OR GRADER ACTIONS

An aspect of practice differentiation appeared in relation to the review and inspection of citations and references within the assessment artefact. These differences ranged from a superficial review to a very close inspection and correction of errors. There was a third approach utilised by lecturers and that was to quickly glance at the reference list but they would do this as a first activity prior to evaluating the artefact. Written assessment artefacts received altered treatment dependent upon which evaluator undertakes the review (Grainger et al. 2008).

#### 7.1.1.h IMPLICIT KNOWLEDGE OR ACTIONS

One of the most difficult aspects identified by interviewees was that of deciding on a final mark or grade for the assessment artefact. Making a judgement could involve the interviewee moving between elements within the criteria to try and arrive at an award which is reflective of their estimation of the value of the assessment artefact. Strategies such as these have been previously highlighted (Baume et al. 2004; Crisp 2010a; Bloxham et al. 2011), where evaluators use a number of information sources, including physical material in the form of the artefact, learning outcomes, assessment criteria and tacit knowledge including individual expectations.

How interviewees interact with the artefact could be at a theoretical level by interviewees asking rhetorical questions or making rhetorical statements. They explore what is presented by talking to the artefact, striking up a one way conversation. These interactions present evidence of ways in which evaluators probe the artefact to develop their sense of its value, all as part of the marking or grading process (Tomas 2013).

Evaluators know of the existence of marking or grading criteria, but choose to interact with them at different points in the process, which was identified in work by Bloxham et al. (2011) and is supportive of earlier reported findings. The existence of marking or grading criteria is not ignored by interviewees,

but they prefer to gain an idea of the artefacts value without referring to it, in the first instance. This outcome confirms findings of Baume et al (2004) where their study participants changed marks on the marking or grading rubric to something which they felt more comfortably reflected their perception of the work. Interviewees 'Jackie' (PA) and 'Toni' (PA) highlight their approach to utilizing the published marking or grading criteria, both seeing the criteria as an adjunct to their initial interpretations.

These findings disagree with research by Tomas (2013), who evidences that marking or grading criteria is explored by their participants prior to any evaluation of the artefact occurring. Interestingly the structure of their research was such that a short initial interview took place, prior to the main data collection interview when expectations were discussed. Briefed participants had more opportunity to perform as expected, including when to discuss the use of published marking or grading criteria. Data in this study has demonstrated a mixed economy in relation to when criteria are referred to, potentially due to no pre PA interview briefing being inclusive of this aspect.

Marking or grading can be portrayed as an objective activity in which decisions are made without recourse to consider the artefact in parallel with the student who produced it. In contrast to this, interviewees in this study, in common with other researchers (Baume et al. 2004) demonstrating a level of emotional labour in connection with feelings of disappointment in the level of achievement, when this was less than expected. This can only occur if the identity of the student is known to the evaluator and could question whether anonymity may have been of benefit (Brennan 2008) (see 6.3.1.a for further relevant findings).

### **Question 3 - What exists to ensure equity or consistency in professional or inter professional marking or grading, this includes inter-rater reliability?**

#### 7.1.1.i OPERATIONAL NECESSITIES

Interviewees to differing levels talked about moderation processes. One perspective was related to a perceived confidence or lack of confidence and trust in the judgements made through the first marking, that a second consideration process had to be instigated. Rather than seeing moderation processes as a way of monitoring quality, 'Jae' (PA) interviewee saw this as a potentially corrective process. A finding which is corroborated by York et. al. (2000). However a further consideration is the potential for the moderation process to proffer an appropriate mark or grade. Previous work has identified a propensity for agreement amongst evaluators when the first mark or grade is known by the second evaluator (Brooks 2012), as would be the case when second consideration is the moderation medium. When a process of blind double marking is employed, with the first evaluators mark or grade not being revealed to the second evaluator differences in marks need to be accepted, with different lecturers valuing different aspects (Brooks 2004).

#### 7.1.1.j TANGIBLE MARKER OR GRADER ACTIONS

Interviewee 'JA' (PA) had a perspective on the utility of annotating an assessment artefact providing evidence to the student that their essay had been thoroughly read. This perspective goes beyond the provision of feedback as a monologue or viewing it as a mode of dialogue between student and evaluator. Neither does this view appear to be related to ensuring alignment of decisions with the assessment criteria in the marking or grading rubrics or achievement of the learning outcomes. It does, as Adie et al. (2013) identify, provide evidence of an approach which takes account of evidence within the artefact to make a judgement on achievement. From this interviewees perspective evidence of annotations across the artefact

presents a view to the student that they have made their judgements from features recognised within the response to the assessment task.

#### 7.1.1.k IMPLICIT KNOWLEDGE OR ACTIONS

A feeling of uncertainty in own marking or grading was thought to be alleviated by a process of second consideration. From 'Toni' (PA) perspective assurance in the mark or grade awarded would come from such a process. This interviewee is concerned enough to voice their anxiety regarding fairness and envisages a moderation process as one which will bolster their marking or grading confidence. The issue not acknowledged, is that unless specifically directed to review the artefact 'Toni' (PA) was concerned about, a moderation process of second consideration may not examine that particular artefact. As a process for quality assurance (Adie et al. 2013; Bloxham et al. 2016), the institution's approach to second consideration, is that it is left to departments to identify how many, and which artefacts are subjected to this process. There is no universally prescriptive approach, and therefore not all artefacts are subjected to moderation. Individual institutions instigate their own second consideration imperatives which could mean, for example, a minimum of 10% artefacts are examined. This being the case, 'Toni' (PA) relying on a process of second consideration to review the artefact they evaluated, and are concerned about, is indeed a long shot.

#### 7.1.1.l MESSINESS OF MARKING OR GRADING

One process for improving accuracy and equity or fairness in marking or grading processes is the inclusion of second consideration, consisting of teams or pairs of evaluators. 'Alex' (CI) highlights their instigation of second consideration and the importance of changing partners to pair with someone they do not know. Reconstitution of moderation teams has the potential to reduce individual assumptions about what one another mean with limited preconceived knowledge of another evaluators approach.

## **Significant Outcomes**

### **- *informing Research Outcomes (RO)***

Research findings have addressed a number of issues already identified by other researchers however findings peculiar to this study are:

- Community of practice - lecturers considered that it was important to know the departmental conventions as well knowing how to execute the task of marking and grading (knowing what to do and possessing 'know-how') **(RO2)**
- Impact of workload on quality of the evaluation and feedback produced when lecturers work with shorter timescales within which to complete the task – **CI (RO2)**

7.1.2 OBJECTIVE 2: IDENTIFY WHAT ARE 'GOOD' MARKING PRACTICES TO DEVELOP A CONCEPT OF 'EXPERTISE' IN MARKING AND GRADING

<b>QUESTION 1: WHAT DO LECTURERS BELIEVE THEY DO WHEN MARKING OR GRADING? (CI)</b>	<b>QUESTION 2: WHAT THOUGHTS OR COGNITIVE PROCESSES OCCUR IN THE LECTURER DURING THE ACT OF MARKING OR GRADING? (PA)</b>	<b>QUESTION 3: WHAT EXISTS TO ENSURE EQUITY OR CONSISTENCY IN PROFESSIONAL OR INTER PROFESSIONAL MARKING OR GRADING, THIS INCLUDES INTER-RATER RELIABILITY?</b>
<b>Parameters of Practice</b> 4.3.2a, 4.3.2c	<b>Operational Necessities</b> 6.3.3a, 6.3.3d	<b>Tangible Marker or Grader Actions</b> 6.3.4a
<b>Knowing</b> 4.3.4c	<b>Tangible Marker or Grader Actions</b> 6.3.4b, 6.3.4d	
	<b>Implicit Knowledge or Actions</b> 6.3.5a, 6.3.5c, 6.3.5d, 6.3.5e	
	<b>'Newbie' Issues</b> 6.3.6a, 6.3.6c	

Table 7 Objective 2

## Question 1 - What do lecturers believe they do when marking or grading? (From CI interviews)

### 7.1.2.a PARAMETERS OF PRACTICE

One interviewee commented on the process of reflection, and the opportunity to participate in a reflective activity, 'Alex' (CI) identified how they had not had a previous opportunity to closely examine their own practice and as such it was something they just did with limited thought about the practice. They considered participating in a process designed to examine evaluator practice was a developmental opportunity, giving time and space to explore a common practice outside the pressure of an assessment period.

How interviewees sought to prepare or induct new staff for their evaluator role was discussed. The usefulness of having a tool around which discussions could be structured was raised by 'Harri' (CI) who used the marking or grading rubric to frame talk about marking or grading.

Interestingly 'Harri' (CI) goes on to say:

*"... they're all practitioners ... got experience in their own right ..."*

This perspective is supportive of the dimension revealed by 'Mel' (CI) and 'Lesli' (CI) to explore the sub theme 'Protecting Clinical Practice' (Chapter 4); the health care academic seems to have a dual role to fulfil using their expert subject knowledge for two purposes. 'Harri' here has identified the second purpose, utilisation from an academic perspective. This interviewees' viewpoint conceives new members of the lecturing team as possessing an interchangeable skills set. This is not the view reported in the literature. McArthur-Rouse (2008) explores the difficulties in transition from expert practitioner to 'novice' lecturer. They identify that there are specific requirements for preparation which should be acknowledged as necessary for those changing roles beyond the provision of handbooks and institution induction days. These authors consider the immense anxiety caused by changes in role from professional to academic, causing confusion, conflict (McArthur-Rouse 2008) and loss of identity and low levels of confidence. This position does not fit with 'Harri's' (CI) assumptions of the skills set of experienced clinical practitioners. The explanation for this position is that

experienced staff have internalised their work role, with knowledge becoming implicit, tacit and underestimating levels of anxiety in staff new to the academy.

Garrow (2009) concluded that induction into local communities of practice for new academics was an important step. Where assessment was concerned, how experienced academics were able to communicate their knowledge of marking or grading processes was deemed to be an important aspect. This fits with the findings of Handley et al. (2013) work which relates to the use of exemplars to facilitate the induction and development of new academic staff into the evaluation of assessment artefacts.

Lastly 'Harri' (CI) questions the lack of continuing professional development opportunities at their institution for developing marking or grading practice, he considers that this is the thing that has the greatest effect on students. Findings in my research demonstrate that interviewees across a variety of institutions feel under prepared for their evaluator role (Trowler and Knight 2000).

#### 7.1.2.b KNOWING

Being familiar with all information pertaining to the assessment artefact was a step in preparing to undertake marking or grading by interviewees. They would have on hand the criteria that students had been given. Two interviewees felt it was important to remind themselves of the exact criteria for evaluation and the exact requirement ensuring use of what Sadler (1989) refers to as *sharp* criteria. These are those which are able to be used for assessment of the artefact, rather than *fuzzy* criteria, those not being overtly called upon for evaluation in this instance.



## **Question 2 - What thoughts or cognitive processes occur in the lecturer during the act of marking or grading? (From PA interviews)**

### 7.1.2.c OPERATIONAL NECESSITIES

There were some good examples of how to articulate application of the marking or grading criteria to both students and in the support of less experienced colleagues. Using excerpts from the artefact itself, and in the case of staff development, demonstrating how the evaluator is reviewing the criteria of 'Application of Knowledge'. Use of exemplars from the artefact is a way of dynamically presenting meaning of the learning outcomes and illuminating the criteria in the real world of marking or grading.

### 7.1.2.d TANGIBLE MARKER OR GRADER ACTIONS

A number of demonstrable actions which a number of interviewees undertook were identified including annotations on the assessment artefact. One interviewee explained that it was primarily to support their development of feedback (Crisp and Johnson 2007). 'Sam' (PA) interviewee made their annotations available to the student.

Whilst Ball et al. (2009) found that there could be issues regarding interpretation and understanding of the annotation, due to the style adopted. These authors concluded that annotations could be a useful practice which could be adopted to support student learning. 'Jo' (PA) identifies annotations as being a good way of delivering specific focused feedback using the artefact as an example by providing a greater number of comments. This interviewee implies that there is insufficient room on the rubric to give detailed comments or advice on certain aspects within the artefact. Opportunity to provide more detailed explanation would enable the language employed to be of a more useful nature to the student. Previous research evidence identifies this as one reason that deters student's use of assessment feedback (Taras 2006; Walker 2009) as the lack of understanding of what the feedback says or indeed how to apply it. The use of on script annotations would be a way of addressing an issue of improving clarity and direct examples of application of feedback.

Where the review of citations or references is concerned, 'Jo' (PA) interviewee gives a careful explanation of their process. This interviewee reviews the accuracy of citations and references, ensuring that they match one another, including appropriate use of punctuation, in line with the institutions published referencing guidelines. These findings are in line with current literature (Sadler 1989; Grainger et al. 2008) and forms part of the judgement of quality of the artefact.

When it came to making notes about the assessment artefact, the majority of interviewees made handwritten notes. One interviewee 'Sam' (PA) used a further document in 'word' to record their notes in an electronic format. This facilitated a speedier approach to the construction of feedback by using the copy and paste function between the word document and the marking or grading rubric.

#### 7.1.2.e IMPLICIT KNOWLEDGE OR ACTION

An approach taken by one interviewee, to immerse themselves in the assessment task to ensuring they were clear about what constituted achievement, was to review previous submissions these included looking at artefacts across the range of achievement. This approach develops knowledge of the assessment task and of previous evaluation decisions, facilitating locating judgement decisions within their local community of practice (Herbert et al. 2014). Reviewing and awarding marks or grades to submitted assessment artefacts following on from this, supports decision making. Evaluators perform what appears to be a seamless act in which they rely upon implied knowledge to complete the marking or grading. However, using artefacts from previous submissions in a preceding academic year firmly bases decision making on underlying rather than purely tacit knowledge.

Interviewees discussed their approach to evaluating the artefact and that it was common to focus on the defects. Evaluators reviewed assessment artefacts to identify all the negative aspects. Instead they were being encouraged by external sources to focus on identifying positive features of the work, giving credit for those. A change in emphasis for the assessment

artefact evaluation will have potential to impact upon both the final mark or grade awarded, and the feedback constructed.

#### 7.1.2.f 'NEWBIE' ISSUES

Those new to a role of evaluating written assessment artefacts identified what would have been most helpful for their development which was in the guise of marking or grading workshops with departmental colleagues. 'Ali' (PA) considered these would be a good way to establish a baseline for the community of practice or department approaches to marking or grading. Workshops such as this can be time consuming but are invaluable methods for inducting 'novice' markers or graders and those new to the assessment artefact for the module or unit of study. For 'Ali' (PA) this approach to support had been well facilitated by marking and grading workshops in a team environment where learning took place through the experience and knowledge of others.

Such an approach may for some new to marking or grading be most effective if introduced at an early stage rather than as for 'Ali' once they had been in the post for over a year. This interviewee then went on to express their thoughts on what would have been helpful this included being given examples of essays from the previous year and examples of feedback. Their proposal was one which saw the inclusion of essays that had been subject to moderation to assure of reliability.

The process for preparation of a new colleague was identified by a 'novice' as one which would have served their developmental needs.

### **Question 3 - What exists to ensure equity or consistency in professional or inter professional marking or grading, this includes inter-rater reliability?**

#### 7.1.2.g TANGIBLE MARKER OR GRADER ACTIONS

Findings in this research have been extracted from a heterogeneous population of health lecturers from 4 higher education institutions. What can be seen is that some interviewees favoured the use of annotations whilst

others did not leading to inequities in the development and presentation of student feedback. My findings thus supports the work of others (Ball et al. 2009; Walker 2009), with the use of annotations as a means of feedback not being used universally. However there had been a move away from artefact annotation related to quality assurance reasons, with evidence of the feedback being taken away by the student. However in the era of electronic marking and feedback, this is potentially an approach that could be explored with a view to improve student engagement with feedback.

## **Significant Outcomes**

### **- *informing Research Outcomes (RO)***

- Identified some utility in the use of CI interview to examine lecturer practice and as a tool for continuing professional development (CPD) **(RO6)**
- HE Health Lecturers taking on a dual role of protecting academic standards through marking and grading, alongside protecting clinical practice from those not fit to practice **(RO3)**

7.1.3 OBJECTIVE 3: EXPLORE NOVICE LECTURERS THOUGHTS ON MARKING OR GRADING

<b>QUESTION 1: WHAT DO LECTURERS BELIEVE THEY DO WHEN MARKING OR GRADING? (CI)</b>	<b>QUESTION 2: WHAT THOUGHTS OR COGNITIVE PROCESSES OCCUR IN THE LECTURER DURING THE ACT OF MARKING OR GRADING? (PA)</b>	<b>QUESTION 3: WHAT EXISTS TO ENSURE EQUITY OR CONSISTENCY IN PROFESSIONAL OR INTER PROFESSIONAL MARKING OR GRADING, THIS INCLUDES INTER-RATER RELIABILITY?</b>
<b>Settling Down to Do the Job</b> 4.3.1d	<b>Student Assessment Literacy</b> 6.3.2a	<b>Operational Necessities</b> 6.3.3f
<b>Parameters of Practice</b> 4.3.2c, 4.3.2d	<b>Operational Necessities</b> 6.3.3b, 6.3.3c, 6.3.3d, 6.3.3e, 6.3.3f	<b>Implicit Knowledge or Actions</b> 6.3.5a
<b>Messiness of Marking or Grading</b> 4.3.3a	<b>Tangible Marker or Grader Actions</b> 6.3.4c, 6.3.4d	<b>Messiness of marking and grading</b> 4.3.3f
	<b>Implicit Knowledge or Actions</b> 6.3.5a, 6.3.5c	
	<b>'Newbie' Issues</b> 6.3.6a, 6.3.6b, 6.3.6c, 6.3.6d	

Table 8 Objective 3

## **Question 1 - What do lecturers believe they do when marking or grading? (From CI interviews)**

### 7.1.3.a SETTLING DOWN TO DO THE JOB

Interviewee 'Mel' (CI) has less than one year's experience as an academic and discloses their process for evaluating the artefact, this requires two readings. Undertaking two readings doubles this individual's workload, but it facilitates them being able to apply criteria in what they judge as a consistent way. This appears to mirror findings from other studies (Sadler 2009b; Bloxham et al. 2011). To make a decision academics will use a three stage process of, read, consider the value of the artefact in light of the criteria being used, and then review the feedback and mark or grade awarded.

Interviewees with significant clinical experience prior to entering the academy voice concerns about feeling ill prepared for their new role of evaluating assessment artefacts. Their introduction often involved entering into a double marking process with their manager, often one of the more experienced and senior academics in the department. 'Mel' (CI) appeared to be expecting more preparation than having their marking reviewed by their manager. It was not evident to this interviewee that this approach could be seen as being taught how to mark and grade, with learning on the job rather than prior to needing to make judgements for summative assessments. One way in which 'novice' or inexperienced evaluators can be inducted is during marking or grading conversations which could occur through moderation discussions with one or more colleagues. Discussions which focus on why and how judgements are made will have a developmental aspect, supporting the development of expertise through an increased exposure to marking or grading decisions (Hoffman 1998). However it is suggested that implementation as well as interpretation of the evaluative criteria is through a socially constructed consensus (Brooks 2012) and this is more easily achieved through conversations with multiple rather than single colleagues. The suggestion raised here may potentially have some utility.

An interesting issue on workload planning was broached by 'Eddi' (CI) identifying how the amount of work and time available can impact upon

marking and grading strategies and ultimately their levels of stress. It appears lecturers can experience a shift in locus of control (Ajzen 2002), with increasing levels of anxiety when they are given more work than they were anticipating.

#### 7.1.3.b MESSINESS OF MARKING OR GRADING

'Mel' (CI) relies on comparing marks or grades they award with those of their manager, thus providing context to their induction in this important aspect of their practice. A boost to their confidence was achieved when the two results were within a few marks, 'Mel' (CI) then felt they were well orientated to evaluating the assessment artefact. This fits with previous literature on induction to marking or grading (Boyd and Bloxham 2014).

'Gerri' (CI) and other experienced evaluators find ways of transmitting their practice to newer colleagues setting out their thoughts on not awarding marks with a 1% or 9% ending, feeling they were too near grade boundaries. This serves to set a benchmark where percentage scales are in use but have a glass ceiling and glass basement within each 10% grade band. Use of such frames of reference can constitute use of normative approaches to constructing marks or grades whilst purporting to use objective criteria. Sadler (2009b) questions whether the mark or grade awarded to the assessment artefact is to be trusted as being commensurate with its value. This appears to be especially true, when the more senior evaluator instigates a glass ceiling on the highest mark or grade they are prepared to award.

Arrival at the final mark to be awarded involves a forward and backward review and readjustment of marks or grades awarded 'Lesli' describes her process for reaching final decisions for a batch of assessment artefacts, reviewing a few and if they find one that they feel is good they look at the others. This approach signifies to adoption of normative referencing, using the one judged as best as a benchmark. However the moving backwards and forwards between and within artefacts represents a patchwork approach to the evaluation, where artefacts are compared to one another to establish achievement rather than being compared to the criteria.

Workload and the institutional approach to providing assessment students with a mark or grade and feedback for an assessment, in shorter and shorter timescales has an impact on those new to the profession. Novice interviewees detailed the number of artefacts they had been expected to mark or grade and the timescale within which this occurred. It seems marking or grading is allocated with no, or limited attention to the experience of the evaluator, including potential to work more slowly as they do not have the underpinning framework of an expert evaluator to rely upon speedy identification of patterns in the artefact (Elander and Hardman 2002; Crisp 2010b). Improvements in costing academic work have seen the arbitrary allocation of time limits on the task of marking or grading. It is unclear whether cognisance of the 'novice' status of evaluators is taken into account in relation to assessment, especially where they assume responsibility for a module or unit of learning to be assessed.

## **Question 2 - What thoughts or cognitive processes occur in the lecturer during the act of marking or grading? (From PA interviews)**

### 7.1.3.c STUDENT ASSESSMENT LITERACY

Where there are markers or graders with less experience, poor academic style used in the assessment artefact may mean that it takes a more detailed or careful consideration to establish its value. 'Maz' (PA) highlighted that they needed to go backwards and forwards through the artefact reading and re-reading. They highlight issues with carrying out the expected evaluation. It appears components which contribute to ease of reading were missing in the assessment artefact, namely style, structure and academic language (Tuck 2015).

### 7.1.3.d OPERATIONAL NECESSITIES

Poor knowledge of the assessment task was a source of increased stress or anxiety for 'Maz' (PA) who felt they had had a limited connection with the module. As a 'novice' marker or grader and one who feels limited connection



with the assessment artefact, they meticulously check each criterion, giving a rationale for selection of elements within it. This interviewee faithfully used criterion referencing to generate the mark or grade to be awarded (Dunn et al. 2002), as limited evaluation experience renders the use of normative referencing almost redundant.

Another aspect raised by an interviewee was where criterion were perceived to be very loosely structured; they struggled with selecting the appropriate categories considering that it was reliant upon their own interpretations. 'Novice' evaluators found this a source of stress, feeling there was limited guidance on how each category on the criteria was weighted, acknowledging that some of their more experienced colleagues appreciated this latitude.

#### 7.1.3.e TANGIBLE MARKER OR GRADER ACTIONS

When faced with the prospect of constructing feedback on the assessment artefact they have evaluated, 'Jae' (PA) adopts a pragmatic approach using elements from the artefact to illustrate what the student had done well. They have found this to be a good way of providing precise feedback signposting with evidence extracted directly from the artefact leaving the student in no doubt as to areas of achievement and why this is judged as so. Developing student's assessment literacy is seen by 'Jae' (PA) as a legitimate enterprise, which should see success in future assessment tasks (see 6.3.2) (Gillett and Hammond 2009; Price et al. 2010; QAA 2016). Utilising examples from their own work and relating that directly to achievement criteria should have the same impact as the approach taken in the study by Smith (2013), which used excerpts from authentic artefacts to explore in groups applying marking or grading criteria. That study established that there were large gains to be made from a small investment in teaching time.

An issue raised by 'Maz' (PA) was the difficulty in being able to distinguish what modifications needed to be made to feedback. They identified the purpose of feedback on a formative or draft work was different to that potentially needed on a summative artefact and this was theoretically in relation to the number of comments provided, with a thought being

expressed as to whether students would actually interact with the summative feedback at all.

This interviewee is seen to be grappling with some of the intricacies and issues related to the marking or grading processes. As a new or 'novice' evaluator it is interesting to note the perceived perspective of the student making limited contact with the feedback, socialisation into the local community of practice, may have cemented such a view (Wenger 2011). 'Maz' (PA) has become part of a group of academics, who through conversations in their department are able to support, share knowledge and practice amongst one another without the need for a formal system.

#### 7.1.3.f IMPLICIT KNOWLEDGE OR ACTIONS

'Maz' (PA) highlights the difficulty of pinpointing the final, despite using a marking or grading rubric. This interviewee appears to see the process as one which is elusive, difficult to grasp. They explain their limited confidence, but participate in conversations with colleagues to increase surety as well as gain support.

#### 7.1.3.g NEWBIE ISSUES

When it came to exploring and understanding the lived experiences of 'novice' lecturers, interviewees in this study were very forthright in expressing their concerns. Interviewees in one way or another articulated a variety of anxieties which appeared to be due to their new role with feelings of being very unprepared and of being an imposter, whilst the remainder of the team seemed to know exactly what they needed to do.

Whilst another interviewee expresses this in another way:

*"... I was ... looking for this magic formula that somebody somewhere would know ..."* 'Ali' (PA);

The interviewees hint that they consider that something concrete will be revealed to them that will be of a great help to their evaluation practice. One of the ways novice evaluators increased their confidence was through continued contact with assessment artefacts throughout the marking or grading period. One interviewee identified another approach to developing

confidence and proficiency which was through producing formative feedback for students during their assessment artefact development, prior to summative assessment.

Some authors identify the difficulty of applying fine grade assessment consistently across modules and academic years suggesting it is difficult to discriminate with precision between 52% and 53% (Bridges et al. 1999). Other interviewees echoed the need for support with 'Jae' (PA) highlighting that despite requests they had not had any workshops. Providing workshops or seminars on the process or techniques used in marking or grading would be appropriate to form a cohesive community of practice within their department or module team.

An interesting point is raised by 'Ali' (PA) regarding evaluating assessment artefacts that discuss clinical practice. They question whether it is possible to assess or evaluate the complexities of clinical practice, exploring the human condition in a tool or rubric by selecting statements which appear appropriate. This is a perspective which has been explored by Trede and Smith (2014).

In spite of an external appearance of confidence in their level of clinical knowledge which they use effectively in their new evaluative role, 'Maz' (PA) at the conclusion of the protocol analysis commented that nothing would be learned from them as they were too new. This provides evidence of an existing dichotomy in this interviewee's academic identity and feeling of self-worth within the academy. Such feelings are explored in-depth by McGregor (2007) using an auto-ethnographic narrative approach to explore their own evolution into an adult educator, highlighting the complexities, anxieties, and realities, as they see it, of academic practice.

The issue of time taken to execute the role is explored by 'Maz' (PA) when explaining how they may take extra time to mark or grade when compared to a more experienced colleague. They considered that they could not quickly read a paper of 4,000 words or more quickly and still feel they had given it the time and attention it merited. They talk of needing time to read through, which is on two occasions to be confident that they can identify the strengths

and weaknesses of the assessment artefact. This two-step approach is confirmed by other researchers (Bloxham et al. 2011; Tomas 2013).

### **Question 3 - What exists to ensure equity or consistency in professional or inter professional marking or grading, this includes inter-rater reliability?**

#### 7.1.3.h OPERATIONAL NECESSITIES

One of the tools used in an effort to ensure equity and consistency in artefact evaluation is the instigation of moderation processes. Novice evaluators reflected upon the process put in place to support their development and ensure that they were marking and grading effectively. 'Mel' (CI) discusses double marking by her line manager, enabling a checking up process to be instigated. Whilst 'Maz' (PA) was positive about having an academic mentor who could ensure that they were marked properly by second or double marking. A process such as this when instigated could ensure a more reliable determination of student achievement as it reduces inconsistencies noted by single evaluators. In the absence of unequal power relations double marking has the capacity to value the variety of opinion which exists between evaluators (Brooks 2004), and uses this as a basis to reach an agreement.

'Maz' (PA) communicates the idea that there was a different treatment of the assessment artefact either due to their status, being a 'novice', or for some other reason. They then identify an approach they would take to increase their own confidence and as a way of confirmation of their competency with evaluation of an assessment artefact, which is to ask for the work to be second considered even when they suspected it was a pass, as a way of bolstering their confidence levels.

#### 7.1.3.i IMPLICIT KNOWLEDGE OR ACTIONS

Interviewees went on to speak about their experiences of moderation processes which are influential in relation to the final mark or grade awarded

to an assessment artefact, with 'Jae' (PA) stating colleagues became upset when their work was second marked.

This interviewee appears to be expressing the anxiety experienced by 'novice' evaluators in the moderation process, especially where power differentials are in evidence. 'Novice' evaluators often feel pressure to defer to the decision of their more experienced colleagues (Wyatt-Smith et al. 2010; Handley et al. 2013) who have evaluation experience along with possessing the tacit knowledge upon which to base judgement decisions. Jawitz (2009) highlights the difficulties faced by new academics in defending a mark they have awarded against that of a more senior colleague the existence of power differentials are an important aspect of marking and grading practice:

#### 7.1.3.j MESSINESS OF MARKING OR GRADING

'Novice' or new markers expressed how they participated in moderation processes or events, feeling marginalised with their perspectives being worth less than their experienced colleagues.

This is evidence of unequal power relations, leaving the 'novice' marker or grader not feeling that their opinion was not valued. In this case the 'novice' marker did not feel they were a full member of the local community of practice. Instead they engaged in '*legitimate peripheral participation*' (Lave and Wenger 1998), as an accepted member acting at a 'Sotto Adagio' tempo, slower and more subdued than their experienced colleagues. 'Mel' (CI) appears to as yet still be on the fringes of their local or departmental community of practice in relation to marking or grading (Garrow and Tawse 2009). 'Mel's' (CI) perception of their position could be related to not having developed sufficient confidence levels in their own practice evaluation and assessment practice, to feel able to challenge others. Second consideration discussions should give rise to opportunities for cohesive *community building* (Bloxham et al. 2016) in an approach to marking or grading that exposes evaluators at all levels to implementation and application of standards and criteria.

Interviewees disclosed a number of issues related to how agreement was achieved between evaluators regarding the mark or grade to be awarded. 'Alex' (CI) explores the tensions experienced when trying to allocate a mark or grade just below the level for a pass or for compensation, being told this was unacceptable. Senior colleagues with more expertise or experience wanting to avoid situations they viewed as contentious and not fitting with the local community of practice (Ecclestone and Swann 1999). Though, this appears to override the need to use the full range of marks or grades available.

## **Significant Outcomes**

### **- *informing Research Outcomes (RO)***

- Experience recast as expertise in marking and grading; being able to perform the task with limited conscious effort **(RO7)**
- Require development of workshops for novice academics to facilitate understanding of components of marking and grading practice including application of relevant criteria or induction into holistic evaluation of assessment artefacts **(RO6)**

7.1.4 OBJECTIVE 4: EXPLORE COGNITIVE PROCESSES AND EXTRANEOUS INFLUENCES ON MARKING OR GRADING PRACTICE

<b>QUESTION 1: WHAT DO LECTURERS BELIEVE THEY DO WHEN MARKING OR GRADING? (CI)</b>	<b>QUESTION 2: WHAT THOUGHTS OR COGNITIVE PROCESSES OCCUR IN THE LECTURER DURING THE ACT OF MARKING OR GRADING? (PA)</b>	<b>QUESTION 3: WHAT EXISTS TO ENSURE EQUITY OR CONSISTENCY IN PROFESSIONAL OR INTER PROFESSIONAL MARKING OR GRADING, THIS INCLUDES INTER-RATER RELIABILITY?</b>
<b>Settling Down to Do the Job</b> 4.3.1c, 4.3.1d, 4.3.1e	<b>Operational Necessities</b> 6.3.3b, 6.3.3c, 6.3.3d, 6.3.3e, 6.3.3f	<b>Environmental Elements</b> 6.3.1a, 6.3.1c
<b>Parameters of Practice</b> 4.3.2a, 4.3.2b, 4.3.2d	<b>Tangible Marker or Grader Actions</b> 6.3.4a, 6.3.4b	<b>Operational Necessities</b> 6.3.3d, 6.3.3f
<b>Messiness of Marking or Grading</b> 4.3.3a, 4.3.3c, 4.3.3d, 4.3.3e	<b>Implicit Knowledge or Actions</b> 6.3.5a, 6.3.5c, 6.3.5d	<b>Messiness of Marking and Grading</b> 4.3.3d
<b>Knowing</b> 4.3.4c, 4.3.4d		

Table 9 Objective 4

## Question 1 - What do lecturers believe they do when marking or grading? (From CI interviews)

### 7.1.4.a SETTLING DOWN TO DO THE JOB

Interviewees explored aspects which facilitated them in completing the task of marking or grading. Interviewees commented on a process to consult supporting documentation prior to producing a mark or grade for the artefact. These documents included the module handbook with learning outcomes as well as the marking and grading criteria. The approach of these interviewees fit with the findings of Tomas (2013).

One interviewee explored an issue which will have implication for assessing an artefact as an isolated piece of work 'Kim' (CI) commented on how difficult it was to evaluate the first piece of work, but then with subsequent artefacts this became easier. A decision on the mark or grade to be awarded can best be made within the context of a batch of artefacts, and thus requires evaluation of more than an isolated artefact. This fits with contemporary evidence on norm referencing and approaches which purport to utilise criterion referencing (Saunders and Davis 1998; Sadler 2009b). This gives rise to potential implications for the granting of extensions where artefacts are more likely to be reviewed alone, making judgement more difficult. When interviewees have explored how they go backwards and forwards within and between assessment artefacts to decide on the mark and grade to be awarded through utilising a process of comparison. This is an unattainable process for a solitary supplementary submission where an extension to the submission date has been granted.

When it comes to deciding on how many artefacts interviewees felt they could interact with, at any one time, 'Kim' (CI) raised an issue which is pertinent when considering lecturer workload which was how they reached saturation by 4 or 5 case studies. The interviewee commented that they stopped '*seeing it*'. This has implications for arriving at a fair mark or grade for the artefact, as evaluators are not able to distinguish either its virtues or shortcomings. Workload therefore needs to be considered in relation to the



ability to construct assessment feedback which could be utilised by students (Higgins et al. 2001; Crisp 2007).

Interviewees then move to reflecting on the criteria or rubrics themselves 'Billie' (CI) explored an issue with applying criteria when one section of the artefact is done very well and another is addressed poorly. They hinted in the end they would just need to take an average to arrive at the mark or grade. This demonstrates the difficulties with pin pointing the exact mark or grade because of selecting areas of the rubric which they feel describes achievement for different aspects of the artefact.

This position agrees with that of Sadler (2009c), who proposes difficulties in the application and use of grading criteria whilst no strong theoretical evidence exists of their effectiveness as tools for evaluating achievement. The potential to apply Biggs' (1996) notion of constructive alignment, to tools to be used to complete evaluations of assessment artefacts is raised. Constructive alignment urges educators to ensure a close association between the learning objectives set for the programme of study, the approach to teaching, student learning and the assessment. To achieve the latter, marking or grading rubrics need to be created to suit the evaluation task by members of the department team, facilitating a community of practice approach (Herbert et al. 2014). A number of these rubrics use the fine grade approach criticised above (Sadler 2009c), whilst taking a more holistic approach to evaluation may be more appropriate. This method would allocate a qualitative description rather than a percentage on the basis of the quality of the assessment artefact rather than marks accrued (Biggs and Tang 2009).

Next the potential for use of different types of criteria for different artefacts is identified by 'Billie' (CI) as important. They reflected on criteria designed and constructed to evaluate an exam artefact and how this could not successfully be used to evaluate a dissertation.

Andrade (2005) asserts that it is appropriate to test the validity of the tool being used to evaluate the assessment artefact. It is evident from my interviewees that there are institutions, departments or module leaders who

devise rubrics for evaluation of individual assessment artefacts, these may suffer from validity and or reliability issues as their suitability for the task may remain untested.

'Mel' (CI) discusses the importance of marking or grading criteria and how, at their institution, criteria are constructed specifically for each assessment, with one of 5 or 6 elements of the rubric being common across all marking or grading criteria. Whereas 'Lesli' (CI) cautions on how criteria can be rendered ineffective due to poor wording of criteria which are needed to facilitate an accurate interpretation of achievement denoted by the assessment artefact (Sadler 2005):

The next area of concern explored by interviewees was in relation to their working environment and its impact on their ability to complete their marking or grading task. Comments from interviewees demonstrate that they found shared office environments not conducive to the level of concentration required, instead they would mark or grade at home.

The subject of working environments in general have been investigated in relation to an increase in sickness absence (Pejtersen et al. 2011), office noise and concentration levels (Banbury and Berry 2005). This latter study is most pertinent to the research reported here. Interviewees highlight the need to be able to concentrate as being an important factor to elicit effective and accurate marking or grading and construction of feedback.

There was also self-awareness amongst interviewees of their best time to complete the task may mean that this could not necessarily be conducive to working in a shared office. Although one interviewee highlighted other potential office distraction such as coffee with colleagues telephone and email. Non-shared environments could be more favourable to developing an atmosphere conducive in creating opportunities to complete assessment artefact evaluation. However, this was only when the highlighted distractions are removed. Thus the environment within which interviewees work within their institutions can prove to be a barrier to them fulfilling their marking or grading roles proficiently.

#### 7.1.4.b PARAMETERS OF PRACTICE

Exploration of this major theme resulted in two interviewees highlighting a sense of dual identity. 'Billie' (CI) positions their need to double check their evaluation of the assessment artefacts mark or grade and feedback as a reflective activity.

Discussing the reflective dimensions of evaluator activity led 'Chris' (CI) to remember their own experiences as a student and the disjuncture between the results gained and their own self-evaluation. 'Chris' (CI) contemplates the impact or ramifications of an unsuccessful assessment submission for students, considering the meaning of this outcome.

For 'Chris' (CI) their reflection is on unsuccessful assessment submissions which involved them in visualising the student's missed career aspirations:

*“... it is really important to them ... that is why I think I can't look into their faces ...”.*

'Chris' (CI) identifies the great importance of the assessment and appears to hint at a feeling of guilt, of being unable to make eye contact with the student (Constanti and Gibbs 2004).

What this interviewee says here resonates with the findings of Smith and Boyd (2012). They reported on health professional academics, mainly nurses, who are expert clinical practitioners, retaining practice and procedural knowledge from their earlier career, which they highly value. Their research was conducted during the induction period of the new academic, which was identified as including any time within five years of their transfer to academia. The need to maintain a live registration with the professional body (NMC, or Health Care Professions Council) was extremely important to interviewees in my study. Interviewees thus maintain a sense of responsibility for maintaining integrity of practice.

This notion of health care academics having a dual role is raised by 'Billie' (CI) and is reiterated by 'Lesli' (CI) who identifies with having professional obligations. This position stems from a requirement to remain on the professional register to fulfil their teaching role. Another point, which is well made by 'Mel'(CI) is the impact academic style including presentation,

structure and organisation of an artefact. Good or excellent writing could result in it masking certain elements within the assessment artefacts which could be evident of poor professional practice. Two interviewees go on to articulate that they consider their responsibilities are to identify patient safety issues as they felt accountable for protecting patients.

The NMC have continued to advocate protection of the public by ensuring all potential new registrants are fit for practice and this remains as a theme in their recent publications (Nursing and Midwifery Council and NMC 2016a; 2016b). A recent study has attempted to investigate responsibilities of HEI in relation to protecting the public through practice education standards and how this function is discharged by education providers (MacLaren et al. 2016). Whilst the fitness to practice requirement addresses the health and character of the student, what is less clear for evaluators is how this is to be addressed within academic work. Where assessment of clinical practice is concerned, the NMC have long advocated continuous assessment, involving a process of evaluating everything that students do (Nursing and Midwifery Council 2009; 2010). Whereas academic assessments of submitted artefacts are evaluated via discrete pieces of evidence either included or omitted which are then used to formulate a judgement.

One interviewee explores the potential connections between the presentation of the artefact and how this is reflected in the approach adopted by the student when they are in clinical practice.

This evaluator equates the pride shown by students in their academic work within an educational setting as being transferrable to clinical practice. 'Chris' (CI) appears to use this as a measure for assessment artefact success revealing the existence of hidden criteria. The difficulty here is that it is a year 1 assessment, and any student at this point in their programme would have limited experience and knowledge of the clinical environment to draw upon. The evaluator knows what a professional approach to practice looks like and therefore perceives that they can identify this within a written assessment artefact. The student on the other hand may have had limited access to professional (health care) practice, along with limited exposure to

higher education academic expectations. When considering this, it appears to be evidence of the evaluator's tacit knowledge, they know what they are looking for but does not state how these expectations are communicated to students.

This begins to highlight typical concerns of health academics when judging work presented for assessment. The difficulties in maintaining a dual perspective are highlighted by 'Terri' (CI) who reflects on evaluating an assessment artefact which presented them with some difficulty. Noting limited academic prowess did not necessarily equate to possession of minimal clinical knowledge. This artefact was being evaluated or judged in two ways. What 'Terri' (CI) does not say is how or if clinical knowledge they required was evident in the task requirements and in the learning outcomes.

'Alex' (CI) an experienced lecturer, being in academia for 19 years plus, adopts the position of an expert practitioner (Berliner 2004). As such they think they are able to execute their role without exerting effort, requiring less conscious cognitive activity while the majority occurs at a subconscious level. Berliner (2004) and Hoffman (1998) when identifying the characteristics of an expert both articulate how they can be distinguished. In summary, experts use pattern recognition to accomplish familiar tasks with speed, this is often done as an automatic response, requiring limited cognitive effort to enable efficient execution. These authors also note that expertise is limited to a particular domain, or in the case of this research and interviewees to a specific subject area or to the activity and process of marking and grading written assessment artefacts.

Learning about and how to mark or grade presents itself as troublesome knowledge, difficult to describe to others and difficult to conceptualise (Meyer and Land 2006). The way 'novice' evaluators in this study have described the low confidence levels and increased anxiety in trying to execute their role is evidence of the troublesome aspects of marking or grading. 'Billie' (CI) raises the possibility of learning to mark or grade as being akin to a journey. Such a journey may not always result in achieving full understanding of the processes, but rather remaining in a liminal state able to operate effectively

but not with the proficiency of an expert. The length of the continuum from 'novice' to expert is considered in the literature on expertise, with Benner's seminal work on 'Novice' to Expert (1984) being of significance to nursing and health care professionals.

A further aspect raised by 'Billie' (CI) is the excellent objectivity of the available rubric at their current institution. They were able to compare this to their experiences of using a rubric at their previous institution which provided limited structure and therefore support for objectively evaluating the artefact. The use of rubrics has been explored by other authors (Price and Rust 1999; Price 2005; Bloxham et al. 2011) who assert the difficulties in construction and consistent utilisation.

#### 7.1.4.c MESSINESS OF MARKING OR GRADING

An experienced evaluator, 'Lesli' (CI) discusses their processes for working with a batch of artefacts. They mark a few and when they found one they felt was of a high standard they go back and review the other artefacts. This provides evidence that normative referencing plays a significant role in the evaluation of assessment artefacts. Indeed Lok et al. (2016) infer that the joint utilisation of criterion and normative reference approaches to artefact evaluation is desirable and could be a way to ensure equity and comparability of standards.

Just how criteria are used to frame evaluator's thoughts and perceptions of an assessment artefact appears to be explored by interviewees. When reflecting on their processes they try to discount the subjective approach to evaluation, despite 'Alex' (CI) feeling these were more accurate than indicated by the marking criteria. The use of objective criteria is advocated (QAA 2012; 2013) to set and maintain standards for assessment in higher education in the UK. Although this is the case, assessment standards and the value placed on assessment artefacts remains open to subjective interpretation and application of differing understanding between evaluators. This is further illustrated by 'Gerri' (CI) who remarks how 2 different evaluators can give the same mark or grade but accompany this with different qualitative comments. Despite the use of marking or grading

rubrics, individual evaluators extract unique evidence from an artefact to illustrate achievement, hence the presence of subjectivity remains. Sadler (2009c) questions the utility of rubrics and suggests their replacement with holistic judgement, as does Yorke (2011).

Introduction of workload models into academic settings to try to implement efficiency savings has had a profound effect upon interviewees practice. Interviewees set their own limits basing this on how much information they can individually process during the time available.

When applying an institutional workload model of 45 minutes per 2000 words a 15000 word dissertation would account for just over 5½ hours of work, falling some way short of the average working day of 7½ hours. Although this does not take into account thinking time, considering the artefact in front of them, weighing up the pros and cons and then awarding a mark or grade and constructing feedback. Such activities are difficult to conceptualise or indeed articulate within a workload model (Soliman and Soliman 1997; Malcolm and Zukas 2009).

the decisions made.

## **Question 2 - What thoughts or cognitive processes occur in the lecturer during the act of marking or grading? (From PA interviews)**

### 7.1.4.d OPERATIONAL NECESSITIES

Interviewees expressed an issue related to their expectations. 'Shirley' (PA) considered the structure of the assessment artefact which was not as they would have liked. There is potential for this perspective to introduce some level of bias or difficulty in the application of professional judgement, which Bloxham et al. (2011) refers to as being a complex task. So whilst the assessment artefact being evaluated by 'Shirley' (PA) had not transgressed the task requirements, it fell short of the interviewees' expectations which may be evident in their final evaluation.

From the interviews conducted it was unusual for interviewees not to have had sight of the marking or grading criteria prior to commencing evaluation of

an artefact. For the (PA) interviews one interviewee was not familiar with the criteria whilst another interviewee had not reviewed the criteria. This is a much fewer number than found in Bloxham et al. (2011) study, where up to 50% of the interviewees had not reviewed the criteria prior to commencing their marking or grading. Those researchers assert that their interviewees had internalised the criteria and they spoke of the artefact in relation to criteria requirements and hence were confident that this was used as a basis for evaluator's judgements.

An interesting observation was made by 'Maz' (CI) in relation to the wording of the criteria and the impact on the evaluation process. The suggestion is that criteria itself can have a greater impact on the grade because of the way it is worded or structured (Silvestri and Oescher 2006; Rezaei and Lovorn 2010). The potential for differences in interpretation are exhibited through acknowledgment of the existence of subjectivity, when individual evaluators arrive at dissimilar decisions on the value of an assessment artefact.

Evidence of this comes from 'Maz' (PA) being advised to award between 20 - 30%, for a poor essay attempt, by different colleagues. A process of second consideration did not eliminate subjectivity. It seems what was required was a process of closer evaluation or being re-marked or re-graded, then a moderation discussion between the two evaluators to reach an appropriate decision (Orr 2007).

#### 7.1.4.e TANGIBLE MARKER OR GRADER ACTIONS

Whilst some interviewees talked of their approaches to review references cited within the assessment artefact it became clear that practise had had to be manipulated. This situation has occurred due to a change in the medium of submission, lecturers at 'Jo's' (PA) institution no longer receive paper copies of written artefacts which facilitated an easier process for the review and checking of cited sources. Interviewees spoke of spending time scrolling from one end of the artefact to another to identify cited sources on the included reference list, therefore a change in submission medium has resulted in a significant change to practise.



However, this is not solely responsible for new approaches to the review of artefact sources. A reduction in time available for the marking or grading process impacted upon methods employed by some interviewees making detailed random searches to check the existence of cited references. Others simply rely on their existing subject knowledge to know if citations are genuine.

In the majority of cases, whether there was a close review or not of all the cited sources, interviewees focused on if they were included according to published guidelines. This is with the exception of one interviewee, 'Toni' (PA) who took a much more lenient view, they felt as it was difficult do they would not penalise the student. Such inconsistencies in the evaluation processes and feedback on the assessment artefact are highlighted which frequently raise concerns amongst students (Carless 2006; Lizzio and Wilson 2008).

#### 7.1.4.f IMPLICIT KNOWLEDGE OR ACTIONS

Review of interview data has revealed an interesting element which has the potential to impact upon marking or grading and award or value placed on an artefact. This occurs when interviewees have differing expectations of two students even though they are being assessed at the same level. An important point here is that the difference in knowledge and or approach by the lecturer has the potential impact upon the final value awarded to the assessment artefact. In this case the lecturer is not following written guidance instead they focus on their implicit or experiential knowledge; knowledge which will not necessarily be held by another evaluator.

Another aspect which seems clear is the process which interviewees went through to decide on the final mark or grade. Interviewees appeared to rely on an intuitive device, one which facilitates the lecturer having a feeling about the mark or grade.

An instinctive impression fostered a level of self-assurance, when their evaluation, utilising the published criteria, was congruent with their initial psychological estimation. Such an approach provides evidence of evaluators accessing tacit knowledge, taking an interpretive view on the artefact. The

process is referred to as 'ephemeral' by Orr and Bloxham (2013) that is the evidence of the process is not visible once the mark or grade and feedback are produced.

### **Question 3 - What exists to ensure equity or consistency in professional or inter professional marking or grading, this includes inter-rater reliability?**

#### 7.1.4.g ENVIRONMENTAL ELEMENTS

A number of interviewees broached the subject of anonymous marking or grading in relation to the difficulties they experienced in executing this effectively. One of the issues raised was in the face of close supervision it was almost impossible to not know the identity of the student. 'Maz' (PA) highlights how easy it is to remember the work you reviewed and thus the identity of the student.

Early work undertaken by Brennan (2008) suggested that there was limited empirical evidence supporting the introduction of anonymous marking or grading. There exists a perception that, where the author of an assessment artefact is known to the evaluator, the presence of bias may negatively or indeed positively affect the mark or grade awarded. Despite a paucity of evidence, the National Union of Students has encouraged students to seek anonymous marking from their institutions via a national campaign (2008). Whereas a more recent attempt at providing an empirical basis to the maintenance of anonymous marking or grading has established that halo bias does exist, as has been suspected by the students (Malouff et al. 2013). Their empirical research presents confirmatory evidence of positive bias in the presence of a positive earlier interaction between the student and evaluator. The same artefact was marked or graded consistently higher where the earlier experience of the student had implied different levels of performance. Anonymous marking or grading does not prevent the ethnicity of the student from being revealed with the evaluator identifying that they could hear the words they were reading being said by the student.

In the presence of an anonymised text it is still possible that the student's ethnic group can be guessed, by writing style alone. However, where the author's name has been seen, as 'Toni' (PA) admits in this case it is much easier to make decision based upon bias. There are particular assessments which possibly cannot be anonymised these include oral presentations, performances or portfolios which contain pertinent individual and or personal reflections, the identity of the student is almost inevitable. However the existence of deficit thinking in relation to ethnic minority students who may not necessarily possess the same social and linguistic capital as their indigenous peers, is a known entity (Yosso 2005). As Garcia and Guerra (2004) highlight, deficit thinking is prevalent permeating society of which lecturers are a part.

Findings from my study have identified benefits to maintaining the current system, with contemporary research evidence available to underpin the notion of lecturer bias where the student is known. Literature has explored the gains that could be made where the veil of anonymous marking is removed the include developing personalised feedback. There is potential to increase student engagement with the assessment task if they consider this will form part of the final judgement on their success or otherwise (Brennan 2008).

Assessment type and the skill of the evaluator to be able to navigate the artefact and produce a mark or grade which accurately signifies its value is an important aspect of the marking or grading nexus. An example of this is the artefact being evaluated which was a complex portfolio containing 3 summative elements. Negotiating these three different components and then identifying how one may inform the other when making a final decision on the marks or grades to be awarded, requires effective assessment practice.

The information obtained from interviewees indicates that in order to ascertain if the assessment artefact has met the requirements of the assessment task they need to determine a number of conditions have been achieved. These conditions relate not only to the academic level, but include appropriate structure, demonstration of knowledge and understanding (see

4.3.4.a), and appropriate language including, if necessary, profession specific (see 6.3.2.a).

Coupled with a concern over workload and time constraints in relation to how soon the results and feedback need to be presented to students and interviewees the amount of time available for review and evaluation felt limited. There was a feeling that they wanted to maintain the same level of attention from artefact number 1 through to the end of the batch of essays.

#### 7.1.4.h OPERATIONAL NECESSITIES

When providing marks or grades for an individual artefact, interviewees highlighted the perspective of feeling constrained by what they were able to award to sit within their community of practice expectations (see 6.3.2.b), 'Maz' (PA) commented on the expectation that a fail would be awarded 35%.

Indeed there is a growing body of literature which explores the concept of communities of practice in higher education, some of which explicitly explores the induction and socialisation of new academics to their work context (Trowler and Knight 2000; Garrow and Tawse 2009). These authors explore the tacit knowledge dimension which underpins learning more than the 'know that', and in my research it can be aligned to knowing what needs to be done to complete an evaluation of an assessment artefact. The argument presented here is that advice given by colleagues regarding the mark or grade awarded is related to the 'know how'. Knowledge of the unspoken rules ('know how') and practices shape the application of the 'know that', unique to that community of practice (Gascoigne and Thornton 2014).

When the process of attaining consistency is explored further as 'Toni' (PA) puts it:

*"... one person's gold is another person's average ..."*

They illustrate how it could be difficult to gain consistency and agreement in the face of such disparate opinions. Where there are two or more markers with limited knowledge of the assessment task, who are inexperienced at applying the marking or grading criteria, this could be problematic. A further adjunct to this would be the possession of limited knowledge of the

assessment standards or approaches of those you are moderating (Lawson and Yorke 2009; Adie et al. 2013).

#### 7.1.4.i MESSINESS OF MARKING OR GRADING

For interviewees completing CI interviews the topic of moderation formed a significant part of their concern with the process of marking or grading. They considered the number of facets which could be part of a process to increase surety in the accuracy of the mark or grade along with confidence of the evaluator. Bloxham (2009) explores four assumptions about moderation procedures which are in place in the UK to ensure the validity, accuracy and standardisation of artefact assessments, the first three only are being considered here. These assumptions relate to accurate marking or grading, review of this via an internal moderation process, engagement of an external system for moderation. The fourth assumption being the final degree award reflects consistency with other degrees awarded at other HE institutions and is considered to be beyond the scope of data gathered within this research study. It is evident that a number of systems and processes have been implemented to ensure effective execution of a vital role and function of a lecturer.

One interviewee 'Billie' (CI) questions the nature and accuracy of marking or grading by a single evaluator when it was simply one individual's judgement. For this interviewee they see a process whereby a single evaluator has sole responsibility for evaluation and award of a mark or grade for an artefact as inappropriate. The excerpt appears to be 'Billie' (CI) questioning that their judgement alone could be what decides on the value attributed to the assessment artefact. This position is reflected in the first assumption examined by Bloxham (2009).

Some interviewees went on to discuss the process of second consideration, viewing this as a failsafe process which will support or verify decision making around marking or grading and is employed when evaluators feel unsure.

Interviewees consider the process of second consideration to potentially increase the validity and accuracy of their mark or grade by gaining agreement. However this verification process is not always achieved as the

potential for disagreement between the first marker or grader and second considerer remains.

An interesting perspective was explored by 'Billie' (CI) in relation to the constitution of teams of evaluators who would be paired on their propensity to be overly generous or meagre in the awarding of marks or grades. This action is aimed at achieving fairness for artefacts being assessed by that pair of evaluators. Institutional processes for the implementation of second consideration means that only a small percentage of the artefacts for assessment are reviewed in this way. The use of second consideration gives an illusion of fairness but in reality this is an altruistic goal that is unachievable. This position supports the second of Bloxham's (2009) identified assumptions, that the presence of internal moderation processes would correct errors in first marking leading to fairness and application of appropriate marking standards, and thus increasing evaluators confidence levels (Bloxham et al. 2016).

Whereas for 'Harri' (CI) they undertake an iterative process of going backwards and forwards over the assessment artefact, establishing confirmatory evidence of none meeting of the learning outcomes. This assures them of the decision made. Interestingly Saunders and Davies (1998) warns of taking too long over the evaluation of an assessment artefact proving in their study that lower grades or marks were often the result. They also warned against re-examining an assessment artefact which has already been subject to careful review against a marking or grading rubric or criteria. However for this interviewee this is their first line of action, and given no fail mark was awarded it is not necessarily an artefact that would be subject to second consideration. Whilst Bloxham's (2009) third assumption regarding moderation suggests that even when internal systems do not reflect expected standards, external moderation processes would ensure that students are assessed against standards which are consistent with those current in the UK university sector. This same situation exists for the artefact which 'Harri' (CI) is reviewing, this may be available for review by an external examiner, but there is no surety that it would be amongst a sample of artefacts reviewed. Both internal and external moderation

systems work on the selection of a sample of available assessment artefacts and as such cannot be the sole guarantor of standards across a range of artefacts, or between institutions (Hannan and Silver 2004; Hudson et al. 2015). For example at one institution, guidance on sampling should include a minimum of 10% of the total assessed work (or 10, whichever is greater) (UoB 2015). Whereas participants in Hannan and Silver's (2004) study into external examining considered that they had discretion over the sample and the volume of work undertaken.

One of the difficulties of the moderation process, whether this is second consideration or double marking is when evaluators may not be reviewing the artefact in a consistent way. This in a sense should be a strength rather than a weakness, which should provide an evidence base to Bloxham (2009) second assumption of the utility and effectiveness of review via an internal moderation process. The correction of errors can be achieved, via internal moderation processes but this is dependent upon the errors being searched for, and the level of experience and knowledge base of the evaluator.

Personal confidence of the evaluator will impact upon their decision making regarding artefact evaluation and therefore their ability to challenge others.

Interviewees highlight the different moderation processes, in common with those in the literature (Grainger et al. 2008; Adie et al. 2013) which can be employed to ensure equity and fairness in the awarding of marks or grades. Each of these procedures has the possibility of arriving at a different decision regarding the mark or grade awarded to the individual assessment artefact. Any decisions are evaluator dependent, inclusive of consideration of their previous marking or grading experience, expertness related to the subject and, knowledge of the assessment task and assessment criteria. Thus evaluation of an assessment artefact is situationally and contextually governed, and bounded by knowledge and experience of the evaluators.

Knowledge or experience of the evaluators may not always be as appropriate as it could be. Interviewees discussed the impact of having had no involvement in delivering module content or setting the assessment task, considering this to be detrimental when they were then involved in marking or grading of artefacts. When this was the case, evaluators in essence

lacked assessment literacy in relation to that single assessment task. This supports work undertaken by Jawitz (2009) who indicates that engaging in various assessment activities give rise to opportunity to absorb implicit understanding which is shared by the collective *habitus* of the local department. Immersion in module teaching or defining the assessment task are ways of developing an implicit understanding of what is required of the assessment artefact and therefore facilitates application of the available criteria for evaluation.

### **Significant Outcomes:**

#### **- *informing Research Outcomes (RO)***

- Individually designed rubrics some question as to their validity and reliability, including their effectiveness at evaluating the artefact type and to do this consistently Marking or grading portrayed as troublesome knowledge **(RO1, RO5)**
- Working environment impact of shared office accommodation in HE institutions and it being detrimental to the concentration required to complete evaluator role **(RO4)**
- Wording on rubric shapes the response to the artefact impacting on the mark or grade awarded **(RO1)**



### 7.1.5 SUMMARY

This chapter has facilitated a review of the findings from each data collection method being brought together. The chapter construction was focused around illuminating each of the four research objectives. As each research method facilitated the uncovering of different types of information and this is illuminated via answering the research questions in relation to the research objectives.

The final chapter will review the research and will explore the findings identified as not previously explored in the literature; the addition to contemporary knowledge is highlighted in Table 10, with a summary of all the research outcomes.

## SUMMARY OF RESEARCH OUTCOMES (RO)

<p><i>Findings which <b>concur</b> with what is known in the contemporary literature</i></p>	<p>Common for evaluators to interact with the artefact twice before finalising the mark or grade to be awarded (Bloxham et al. 2011; Tomas 2014)</p>
	<p>Evaluators find it hard to discriminate between fine grade differences in achievement which are exemplified by the allocation of marks or grades e.g. 63% or 66% (Yorke et al. 2000; Yorke 2010; Sadler 2014)</p>
	<p>Assessment work best completed within the context of responding to a batch of artefacts for evaluation. Much better not to complete isolated evaluation tasks due to normative referencing that was identified as part of approach even when applying marking or grading criteria - (Saunders and Davis 1998; Sadler 2009b)</p>
	<p>Novice markers and graders feel under prepared for their evaluator role (Trowler and Knight 2000)</p>
	<p>Normative referencing was used by evaluators to aid judgement (Hand and Clewes 2000; Sadler 2009b; Lok et al. 2016)</p>
	<p>Structure and wording of the marking or grading rubric exert a great impact on the achievement of the assessment artefact (Silvestri and Oescher 2006; Rezaei and Lovorn 2010).</p>

	Evaluators used criteria post marking or grading event - (Bloxham et al. 2011)
<i>Findings which <b>contradict</b> what is known in the contemporary literature</i>	Evaluators may apply criteria post decision about mark or grade as a way of articulating and justifying it - not evident in Tomas (2014)
<i>Findings <b>not previously considered</b> in the contemporary literature</i>	<b>RO1.</b> The 'rubric paradox' evidenced by the acceptance and ready utilisation of rubrics by novice and experienced evaluators yet criticised in the literature. Alongside identification of the challenges associated with producing valid and reliable criteria/rubrics for a range of assessment artefacts whether these are generic or bespoke.
	<b>RO2.</b> Community of Practice to be developed amongst marking or grading teams in a more overt way, taking part in marking or grading conversations pre assessment and post assessment moderation events. Adoption of this approach as a way of improving consistency in marking and grading within a department and supporting the development of novice evaluators.
	<b>RO3.</b> Dual identity of health lecturer, related to upholding standards and fitness to practice of students. Evaluators use knowledge from professional healthcare practice as a benchmark when evaluating assessment artefacts. These tacit standards may not be evident in the assessment task objectives or learning outcomes.
	<b>RO4.</b> Shared working office environment and its impact on being able to complete the evaluative task
	<b>RO5.</b> Messiness of marking and grading practice and troublesome knowledge.

	<p><b>RO6.</b> Interview - PA or CI as a way of examining marking or grading practice - could contribute to Continuing Professional Development, moderation and 2<sup>nd</sup> consideration to illuminate assessment practice and how evaluation decisions are reached to develop novice evaluators.</p>
	<p><b>RO7.</b> Experience recast as expertise in marking and grading; being able to perform the marking or grading task with limited conscious effort embodies unconscious competence. Evaluators can complete marking and grading without mindful attention.</p>

**Table 10 Summary of research outcomes**

# Chapter 8 CREATIVE SOLUTIONS

*“The man who has no imagination has no wings.”*

Mohammed Ali

This final chapter brings together the main findings from the research whilst exploring my contribution to contemporary knowledge and debate in relation to marking and grading by lecturers in HE. Firstly there is a return to the philosophical underpinning and methodology for completing the research which illuminates the approach to data collection and therefore the resultant findings. A critique of data collection methods is included to highlight any advantages and disadvantages encountered. The chapter then goes on to explore the research study limitations and researcher reflexivity before settling upon a review of the main findings. Outcomes of the research form the basis for the exploration of findings in this chapter, these being findings that have not previously been considered in the contemporary literature in this context on marking and grading (Chapter 7). These have been generated by carefully relating the data to the study objectives and to answering the three set research questions. Finally I present a model for exploration and utilisation in the induction of ‘novice’ evaluators, to more quickly support their introduction to and continued development in marking and grading practice. Recommendations for further research are drawn from findings here, but also with a wish to explore the use of the model for supporting marking and grading practice development.

## 8.1 MODEST BEGINNINGS

*Nobody can go back and start a new beginning, but anyone can start today and make a new ending*

Maria Robinson

The hermeneutic phenomenology of Heidegger (Mackey 2005) was utilised as a vehicle with which to explore the everyday experiences of lecturers in HE undertaking marking and grading of assessment artefacts. Taking this philosophical approach to data collection allowed for consideration of in depth explanations and descriptions of the lived experiences of interviewees.

Heideggerian phenomenology encourages exploration of interviewees' narratives to increase or achieve understanding of their phenomena, which is the focus of the research. The selected data collection methods sit well within a methodology which attempts to explicate the meanings given to the acts of marking and grading and all associated with these phenomena.

Two types of interview method were purposefully selected to enable a close exploration of current marking and grading practices. Each method led to the retrieval of a large amount of in-depth and complex data requiring consideration and interpretation. Heidegger states there is no understanding without interpretation (Koch 1995) and yet in order to facilitate understanding the interviewee has to be prepared to share their story. There was an acceptance of storytelling as a legitimate way for interviewees to begin to be immersed in the world of the interviewee, gaining insights that would not normally be possible. Marking and grading is traditionally a practice undertaken in isolation, with little being known about the cognitive thoughts and processes which accompany such a complex decision making activity. Cognitive Interviewing interviews took place post marking and grading events. Interviewees were instructed to reflect on a particular assessment experience focusing on a specific artefact. Interviewees were able to produce a coherent story depicting their approach, processes and in some cases understanding of their role in evaluating an assessment artefact. Holstein and Gubrium (1995) talk of the active role of the interviewee in constructing the story of their phenomenon of which the interviewer wants to achieve closer contact with. According to those authors, interviewees enter the process of an interview with their own interpretation of the event, including individual motivations for releasing or indeed suppressing information from becoming part of the data set. This means interviewees are active agents in the information gathering process, gatekeepers to a phenomena of which the interviewer wishes to gain understanding and insight.

Protocol Analysis interviews allowed access to the world of marking and grading vicariously, as interviewees were instructed to verbalise all thoughts

occurring during an assessment artefact evaluation. A post event interview took place to clarify any details of the approach to artefact evaluation. Again Holstein and Gubrium's (1995) position highlights that interviewees will possess varied intentions for the level of disclosure during this portion of the data collection process. However interviewees have less control over what they include or withhold during a PA interview, as the method calls for the verbalisation of all thoughts and processes as they come into being, in the mind of the interviewee. An authentic story of marking and grading practice is thus retrieved, with some deeper exploration offered post PA interview during debrief.

A desire to uncover the story of marking and grading from interviewees in an authentic way, facilitating the exploration of their experiences which are bounded by contextual, situational and cultural aspects has been achieved. As a result of sustained attention on the philosophical perspective to enquiry, the methodological approach and method, a close fit has been accomplished (Mackey 2005).

## 8.2 REFLEXIVITY

***"a person of a growth mindset has a potential that is unknown (unknowable)  
... everyone can change and grow through application and experience and  
that it is impossible to foresee what can be accomplished with years  
of passion, toil and training"***

*Dweck 2006 - page 7*

The concept of reflexivity has been addressed in Chapter 2 and is also referred to in Chapter 5. My role as researcher has been fundamental in shaping this project. At the outset I identified that I did not approach this topic bias free, with little or no knowledge. Previous experiences have been instrumental in facilitating working with and analysing the data. At times during this thesis I have used the first person to indicate my influence on the this work (Davies and Hughes 2014)

*"The person who has gathered the qualitative research material is uniquely you. All data is filtered through your intellectual self". (p 266)*

At the outset I state my role as that of a story teller, being able to co-construct the narrative around the lived world of lecturers in HE completing marking and grading. Engaging in hermeneutic phenomenology allowed me to gain a view of the world of other health professional academics evaluating written assessment artefacts. Being part of that community of practice has afforded me different insights and the ability to make interpretations which are situated and contextualised in ways unavailable to an outsider. Instead of this being viewed as a limitation I have seen this as an opportunity to come to a nuanced understanding, reaching the goal of this thesis, demystifying marking and grading processes.

A further concept which is relevant to the process of reflexivity is to think about the issue of power and from a Foucauldian perspective (Allen 2002) the impact of this on the research output, in the guise of data. Within my own institution and indeed own faculty I was well known to be connected with marking and grading and as such this may have had two outcomes to either encourage or discourage recruitment of evaluators to my study. As lecturers were approached via faculty wide advertisement on the intranet they were free from coercion and self-selected to participate. Conversely it is acknowledged that a perception of my increased knowledge base in relation to the topic may have had a detrimental impact upon recruitment. New academics or those with experience (but remained unsure in their marking or grading practice) may have been more reluctant to volunteer as they may have perceived me to be an expert in this field. In the end this did not seem to be the case, with 15 lecturers from my own faculty volunteering to participate. The processes for recruitment at external institutions is explored in Chapter 2, however, the effects of being known as a 'marking and grading expert' should have been lessened. I did not feel this to be the case, interviewees had drawn conclusions about me prior to our meeting, however, my experience in interviewing, in many different contexts was put to good use, in establishing an atmosphere where shared story-telling could be entered into. Interviewees commented on the process, with this data being retrieved at the end of both interview methods and coded as 'Outliers':



*“ I think you’re kind of patient way makes it like very easy to explore it out loud because it just stops feeling like it is a strange thing because you are just kind of encouraging ... so then I think I wouldn’t have been able to say those things if I hadn’t felt quite so protected” ‘Ali’ (PA)*

*“I thought I would find it more, I thought I would be more anxious, I am sure it’s partly down to your manner” ‘Jae’ (PA)*

A further insight which came to light, at the end of the data analysis process was in relation to the seeming absence of acknowledgement of moderation processes by PA interviewees. At commencement of each of their interviews they were asked to ensure the artefact they were evaluating was subjected to second consideration as an approach to reduce potential detrimental impact of this technique. On reflection this reduced the opportunity for a naturalistic exploration of this theme during the PA interview itself.

## 8.2.1 THESIS CHAPTER SUMMARY

Chapter 1 presented a focused review of the literature, with seven discrete areas selected for closer attention from the vast array of literature on assessment. The choice reflected a conscious decision to ensure closer examination of areas considered to be of most significance to the subject of this thesis. The literature focused upon the following topic areas: marking and grading criteria or rubrics, normative and or criterion referencing, completing marking and grading as an activity, the assessment literacy of staff and students, relevance of communities of practice for achieving marking and grading or written assessment artefacts, tacit beliefs in relation to marking and grading and expertness and expert practice. As my thoughts and understanding of the subject became honed I was able to confine the literature search to elements to unambiguously underpin the research. In essence a hermeneutic approach was taken to retrieve literature (Boell and Cecez-Kecmanovic 2014), iterative in the sense of moving backward and forwards, in and out of searching, retrieving and reading, with further searching for information which would deepen understanding.

Chapter 2 was then focused on exploring the philosophical and methodological underpinning of my research. An examination of my understanding of epistemology and ontology (2.1), led to gaining a good grasp of what I feel constitutes knowledge and how this is derived. This manifested itself into the methodological approach for gaining information to best respond to the aims and objectives established for this research enquiry. A set of three research questions were then devised to ensure that the enquiry remained focused on specific elements of the marking and grading puzzle, and remained concerned with responding to the stated research aims. The methodological approach which most closely aligned to my philosophical stance around knowledge, its existence and construction was that of hermeneutic phenomenology (2.5.2). My desire was to hear the lived experiences of lecturers in HE engaging in the activity of marking and grading, uncovering their thoughts on their evaluations of written assessment artefacts. I considered that a face to face interaction was the best approach, all the while acquiescing to the tenet of active interviewing (2.5.3). This emerges in an acceptance that interviewees have their own part to play in shaping the data gathered, by their responses or their withholding of information. Despite this there still remained this desire to hear the truth of lived experiences of lecturers in higher education as perceived by the 26 interviewees who were recruited to my study. It was important to give attention to my privileged position as a practitioner researcher (2.5.7), holding some understanding of the historical, contextual and sociocultural positions of health lecturers as a member of this wider community.

Modified framework analysis (2.8) was adopted as the approach for interaction with the data. This appeared to be flexible enough to use for analysis of data from the two research methods selected for data gathering from my interviewees. The selection of this method for analysing the data was a pragmatic decision, facilitating synthesis of findings and the extraction of key contributions to knowledge on marking and grading. The benefits of using this method was the systematic treatment of both data sets following the seven steps in the framework analysis method. The drawback identified

fairly early on, was the non- standard way in which I had conducted the interviews precluded the interview questions being used to develop an analytical framework. Instead a hermeneutic approach ensued where through listening to the interviews and reviewing the transcripts common themes were identified and then used as the framework for presenting the findings which emerged from the data. This modified approach gave a good fit with the philosophical underpinnings of my research.

Data for my research was gathered by two different methods and these are presented in detail in Chapters 3 and 5. Both methods facilitated face to face connections with interviewees enabling faithful adherence to a hermeneutic approach, the desire to hear from lecturers their authentic stories on the subject of marking and grading written assessment artefacts. The first interview method is presented in Chapter 3, the Cognitive Interviewing which was found to be a novel application, as use with academics has not previously been reported in the literature (3.1). I developed an interview protocol to identify specific probing prompts to aid the gathering of data of sufficient depth and pertinence (Appendix 6). The second interview method, Protocol Analysis (Chapter 5), has a history of application with lecturers in higher education, with proven validity and reliability as a data gathering tool (Crisp 2008; Boyd et al. 2009; Bloxham et al. 2011; Orr and Bloxham 2013).

The findings Chapters 4 and 6 are set out under identified themes with information gathered from interviewees. These findings are then summarised and synthesised in Chapter 7, each of the research objectives used to compare and contrast findings in relation to responses retrieved for each of the research questions. This led to an intricate combination of responses, maintaining sight of the research objective that they were responding to. Completing this synthesis facilitated the compilation of a summary of research outcomes (Table 10). The findings not previously considered in the literature are being explored here (8.4, 8.5), with some further discussion of their implications. The culmination of this research has been the development of a set of models, detailing components of the marking and grading practice to more quickly support newcomers to develop

expert marking and grading practises. A unique combination of research methodology and methods, research participants and, researcher approach has facilitated the surfacing of a new approach which can be explored with lecturers in HE. This original combination has resulted in new perspectives on a long standing problem.

### 8.3 LIMITATIONS

*“He who is not courageous enough to take risks will accomplish nothing in life.”*

*Mohammed Ali*

Interviewees were all undertaking an authentic assessment activity and there is a theoretical potential that the PA data collection method could have an impact upon the way marking and grading was performed has been identified (See 5.2.3). Therefore it was important to minimise any effects upon how the artefact was evaluated and the final mark or grade awarded. All 11 PA interviewees were asked to ensure second consideration occurred for the artefact that had been used for data collection. This instruction to interviewees in hindsight impacted upon the spontaneity of interviewees in discussing their approach to moderation of assessment artefacts. There was no way of identifying what systems these interviewees would have adopted for review of any work they had evaluated. However interviewees involved in CI interviews, of which there were evaluators from each of my data collection sites, the majority discussed moderation processes. An inference could be made that PA interviewees, from these same institutions would not have overlooked this component had a prior suggestion not been in place.

Cognitive Interviewing was an interview method which has not been previously used in HE to explore marking and grading processes as such it had not been tried and tested in relation to validity and reliability. The technique is predominantly researcher led guiding the interviewee in a deep post task reflection. The aim of the technique is to improve recall along with accuracy of past events. Recall and memory have been shown to be fallible

but the application of specific techniques along with training in their use can increase the accuracy with which events can be remembered and verbalised.

A possible limitation with both interview techniques was the potential to uncover poor or unsafe practice identified by interviewees during their evaluation of assessment artefacts. To prevent interviewees from feeling insecure about verbalising any such identification, ground rules were set out at the beginning of the interview, where I was able to reiterate my approach as a registrant with my professional body. One such case occurred, identified by the interviewee with a fail mark or grade being subsequently awarded to the assessment artefact.

## 8.4 MAIN FINDINGS

The overall aim of this research was to investigate the process of marking and grading from the perspective of the lecturers completing assessment artefact evaluations. On completion of this investigation the hope was to recognise good marking and grading practice deriving a model of 'expert' practice aiding the induction, development and proficiency of novice markers. A set of models for exploring and developing expert marking and grading practice have been identified for utilisation in new staff induction and continuing practice development opportunities. Thus the aim of examining these processes, from the perspective of interviewees, was achieved by utilising a phenomenological approach employing two data collection methods. Chapter 7 presented a systematic approach to the data in relation to each of the four research objectives. This ensured data from each of the major themes could be linked with the research objective, and the particular research question to which it was addressed.

### 8.4.1 "VENEER OF GOODNESS"

Written assessment artefacts are presented as ready for evaluation in many ways. When there are a large number of artefacts requiring evaluation, lecturers can quickly lose heart when there seem to be none which have

followed the guidance set, appear to have little connection to the task or are written with what appears to be a poor command of English. Interviewees in my research have corroborated findings from other studies (Greasley and Cassidy 2009; Bloxham et al. 2011) that the presentation of the work has a huge impact upon the marking and grading and therefore final evaluation. Artefacts that are poorly written with many spelling and grammatical errors are very difficult to read. Often this is accompanied by poor structure which goes beyond individual sentence construction, but includes limited understanding on the appropriate use of paragraphs with writing being characterised by having too few or too many. Then there is the artefact that appears once in a while which does all of the above, perfectly, and thus, states 'Kim' (CI), is "... *a joy to read* ...". Often interviewees referred to this in relation to the way an artefact was written and expressed, showing few errors and having a good flow, selecting the right words to convey meaning eloquently. Good academic style influences lecturer's psychological approach and final evaluation of the assessment artefact.

However all assessment artefacts require thorough reading to get beyond the surface impression. This should mean that an artefact which has been well written with good academic style including perfect spelling, grammar and syntax is judged appropriately and that evaluators are not overly reliant upon such cues to anchor their judgement. Such an artefact may belie content which is limited in depth and breadth or be of limited relevance to achievement of the learning outcomes for the specific task or module. 'Billie' (CI) refers to this as a "... *veneer of goodness* ...". From my research findings it seems that this is something very experienced evaluators need to guard against, more so than their novice counterparts. Those new to the role appeared to be experiencing greater anxiety around their ability to complete the marking and grading function correctly. Thus they more carefully review any set learning outcomes and the criteria against which the artefact being judged, using this information to carry out that judgement and not tacit knowledge retained from a previous iteration. Experienced evaluators may operate in a mode which equates to conscious competence (Figure 1), a

state of proficiency, journeyman or competence in relation to the skill acquisition models (Table 2). Whilst there are others who will progress towards unconscious competence also described as a state of expertise, or proficiency (Table 2).

The use of rubrics was popularised in the mid 1990's, but have been under increasing fire regarding their robustness (Jonsson and Svingby 2007; Sadler 2009c; Yorke 2011). Their use has increased as they are thought to facilitate an increase in objectivity when reviewing and thus reaching a judgement on the value of any artefact presented for assessment. Literature exists regarding concern for students to develop their assessment literacy and thereby improving their chances of succeeding at an assessment, with a few authors having carried out extensive work in this area. Students are introduced to the criteria by providing an opportunity to discuss them, developing understanding of their application in the evaluation of an assessment artefact (Rust et al. 2003; O'Donovan et al. 2004; 2008). The necessity for students to become more involved in an appreciation of the marking or grading criteria or rubric is also voiced by Sambell et al. (2012) and this has to consist of more than showing students the rubrics or criteria. Instead students need to be involved an authentic opportunities to apply the criteria to their work, developing their self-assessment skills. Opportunities to design in peer assessment activities into the learning activity will enhance the chance for practising such skills. Improvements in the depth of student appreciation of marking and grading rubrics was achieved by providing examples of previous authentic assessment submissions for review (Wimshurst and Manning 2012). Students were required to engage in an activity, utilising the rubrics, constructing feedback and awarding a mark or grade to a specified number of artefacts. Use of genuine examples of past submissions supports students in understanding what their finished work should look like, facilitating explanation of the assessment criteria and learning outcomes coming alive in the artefact they are reviewing.

Using appropriate criteria for evaluating the artefact - interviewees from both methods highlighted issues with using unfamiliar criteria and using criteria

which did not seem to facilitate the judging of particular components in the artefact. Utilising marking and grading rubrics which have been developed for the particular assessment type appears to be important in facilitating more accurate judgement of the assessment artefact. 'Billie' (CI) talked of the tool which was developed for an essay examination being used to evaluate a research protocol. This application proved to be ineffective as there were sections on the essay examination marking and grading rubric which either did not fit or indeed were missing. Using an inappropriate tool makes an effective and objective judgement more difficult if not impossible to complete. Another issue broached by one interviewee was the possibility that they were unfamiliar with the marking and grading tool. 'Toni' a PA interviewee referred to this in their interview "... *it's not a marking grid that I've seen before ...*". As an evaluator it appears to be important to know the marking and grading tool to be applied to the assessment artefact. This knowledge facilitates being able to identify areas of achievement more easily.

However Yorke (2011) maintains that the measurement of achievement via evaluation of an assessment artefact is tainted by a multitude of factors. These include institutional procedures set to govern assessment of which moderation is an example; issues relating to the tools themselves and the ability of evaluators to use them consistently to discriminate between fine grade descriptors which are either stated numerically or qualitatively; and the capabilities of individual or teams of evaluators, including accounting for individual bias, to make a fair and accurate judgement of the merits of an assessment artefact. These difficulties have been previously been purported to exist (Sadler 2013; Bloxham et al. 2015; Bloxham et al. 2016).



### **Research Outcome 1:**

**RO1:** The 'rubric paradox' evidenced by the acceptance and ready utilisation of rubrics by novice and experienced evaluators yet criticised in the literature. Alongside identification of the challenges associated with producing valid and reliable criteria/rubrics for a range of assessment artefacts whether these are generic or bespoke.

Rubrics were structured in different ways at different institutions with the lowest mark or grade on the left side of a matrix moving, to highest on the right at one institution to being the opposite at another. The names for components within the rubric also differ between different departments and between institutions, it is therefore important that evaluators make themselves familiar with the composition of the marking and grading rubrics in their locality. Despite these aesthetic differences the underpinning elements of evaluation remained the same at the four institutions where I gathered data, with note being taken of the presentation of the artefact, use of references and literature, content, and analysis. These appeared to be treated as universal criteria that is, those features which appear to be necessary elements of each artefacts presented for academic assessment. These components of the artefact required acknowledgement of their existence or otherwise and then the award of a mark or grade for each. The major differences exist through what these components are named and the value placed on each of them.

The majority of interviewees in my study did not question the use of marking or grading rubrics for evaluation of assessment artefacts, they were therefore accepted as part of the necessary marking and grading nexus. Interviewees instead talked of the support they gained from application of a marking or grading rubric when making decisions about achievement. Interviewees also broached the topic of consistency viewing the utilisation of rubrics acting as a guide to evaluators to ensure that, amongst a team, they reviewed the same elements in an assessment artefact. Interviewees suggest that rubric use as being a good tool to facilitate discussion amongst evaluators. However both Sadler (2009c) and Yorke (2011) are dismissive of the ability of rubrics to

produce benefits to the evaluation of assessment artefacts, with limited evidence of effectiveness or indeed encouraging agreement between evaluators. Nevertheless there is an acknowledgement that locally designed marking or grading rubrics could illuminate the tacit knowledge held within teams or departments thereby improving understanding and application during assessment artefact evaluation (Dawson 2017). Baume et al (2004) identified that despite the presence of rubrics lecturers still relied upon intuitive understandings to make judgements changing marks or grades, arrived at through rubric application, to match internalised levels of achievement. This illustrates Dawson's (2017) perspective who emphasises the complex nature of making judgements utilising evaluative rubrics and the necessary expertise that this requires.

A further issue worthy of further consideration is the practice of developing different marking and grading rubrics for different assessments, at each iteration. In some institutions this was a common occurrence as module leaders constructed an evaluation tool for their module assessment. Validity and reliability of this new tool could be questioned. As it is new and not previously used there is no way of knowing whether it will measure what it set out to measure or do so consistently and in every circumstance. The development and application of bespoke marking and grading criteria or rubrics increases the risk for reliability issues. Potentially an argument could be constructed for increased validity of such a tool as it should be closely related to the learning outcomes for a specific unit or module of learning. Any tool to be used for evaluation requires testing for reliability and validity (Reddy and Andrade 2010). Student constructed marking or grading rubrics could also suffer from validity and reliability issues, for the same reasons as cited above (Orsmond et al. 2000). One proposed advantage of this would be to develop student assessment literacy. However as alluded to earlier simply showing criteria or rubrics to students does nothing to engender understanding, additional input is required. For Gibbs and Simpson (2004-05) the design of the tool for marking and grading is of

limited relevance without engagement in accompanying conversations ensuring students and staff develop a common understanding.

### **Research Outcome 2**

**RO2:** Community of Practice to be developed amongst marking or grading teams in a more overt way, taking part in marking or grading conversations pre assessment and post assessment moderation events. Adoption of this approach as a way of improving consistency in marking and grading within a department and supporting the development of novice evaluators.

There is a suggestion that developing a community of practice may be a preferred route to cultivate and maintain a common approach to marking and grading practice (Price 2005). Communities of practice (CoP) have been seen to support the development of a cohesive approach to reaching and agreeing marking and grading decisions. This was demonstrated in this research by early career evaluators being supported by an individual, usually a colleague with more practice experience, but not necessarily portraying exemplary practice or expertise (Berliner 2004). Further to this, participation in department moderation events and assessment committees will facilitate discussion, sharing practice ideas. This could occur by module team organising discussions between lecturers' involved in marking or grading of the assessment artefact, as a group, prior to the evaluation period commencing. Such conversations contribute to developing common understanding and application of the tools used for evaluation to support development of a local CoP improving inter-rater reliability (Jonsson and Svingby 2007). The development and application of rubrics within a community of practice would increase the validity and reliability of such tools over time. However taking a more strategic view could result in developing a cohesive framework within which to convene and engage in artefact evaluation conversations. At this time clarity on definition and expectation of the articulation of learning outcomes could be provided with a discussion to ensure shared understanding, and how this will then be interpreted through application of holistic or analytical criteria in an evaluation (Sadler 2009d). All consistent with a desire to maintain or develop a CoP approach to

marking and grading resulting in a possible reduction in any questions about reliability. Similar discussions will also take place at the end of the marking or grading period, prior to release of feedback to students, but would have to be built into the timescales set between submission of artefacts and expected availability of results. As one interviewee, Jae, commented “... *you can't have it both ways ...*” referring to the quality of assessment and the continued assault on time available to complete the task. Meeting student demand for reduction in waiting time for marks or grades continues to erode the possibility of implementing strategies to design in shared understanding and agreement on achievement standards via an overt approach to developing communities of marking and grading practice.

#### 8.4.2 “THAT IS WHY I JUST CAN'T LOOK INTO THEIR FACES”

Evaluators were faced with the dilemma of assessment artefacts which do not warrant being awarded a pass mark, with this then having implications regarding students' being able to pursue their chosen career. One interviewee 'Chris' (CI) made a striking comment regarding a re-submitted artefact which had, for a second time, been unsuccessful at achieving a pass mark “... *I can't look into their faces ...*”. This demonstrates the internalised responsibility that can be experienced by evaluators when the author of the unsuccessful artefact is known to them.

#### **Research Outcome 3**

**RO3:** Dual identity of health lecturer, related to upholding standards and fitness to practice of students. Evaluators use knowledge from professional healthcare practice as a benchmark when evaluating assessment artefacts. These tacit standards may not be evident in the assessment task objectives or learning outcomes.

Health academics made up the entire sample of this research cohort. However it did remain heterogeneous as a range of health professions were represented. Evaluators were concerned as to whether what students present for evaluation was a true representation of what would occur in the clinical practice environment. Evaluators in my research remained unsure of whether the written assessment artefact was the appropriate way to assess

knowledge and its application by the student in their clinical practice role. Rather than questioning the assessment criteria or rubric lecturers appeared to be interrogating the appropriateness of testing transfer of knowledge to practice via the previously identified barrier of the theory practice gap (Monaghan 2015).

This begins to raise the issue of a perceived dual role of lecturers. They participate as part of the academy whilst continuing to protect clinical practice, cognisant of human factors theory (Carthey and Clarke 2010; Bromley 2011), and the part individuals play in catastrophic errors in patient care. Lecturers therefore appear to seek reassurance from the assessment artefact that students would be able to put the theory they express into practice, and to do so effectively, reducing patient risk of being exposed to errors. There seems to be evidence from interviewees that they use tacit knowledge from their field of healthcare practice as a yardstick against which to measure the assessment artefact. Whilst these standards may play a limited part in published learning outcomes or marking and grading rubrics they nevertheless are an informal part of the requirement for achievement. This illustrates the subjective nature of assessment as different academics with different levels of expertise or areas of practice may have a different set of measures which they will apply to the evaluation situation.

#### **Research Outcome 4**

**RO4.** Shared working office environment and its impact on being able to complete the evaluative task

Interviewees from both interview methods explored issues related to their working environments. The main issue was related to the increase in shared office space, which diminishes the ability to concentrate. More people often means an increase in the level of noise with extra computer equipment, telephones ringing, other members of staff from different departments wandering in to find a member of staff who, is invariably not available. Two interviewees raise issues with their working environments where they desire:

*“... a quiet place so I can concentrate ...” Kim (CI);*

*“... do it where you are unlikely to get disturbed ...” Maz (PA).*

They both highlight the difficulties of needing to complete a task which requires a particular level of attention and deliberation. It appears to have become accepted practice for academics to work at home, however this does not present the opportunity to maintain a separation between home and work, impacting upon work-life balance (Kinman and Jones 2008). Along with this, in an era of electronic assessment lecturers may not have the equipment to carry out work at home, complying with contemporary health and safety criteria related to work station assessments (Sonne et al. 2012). However it seems PA and CI interview findings were congruent in identifying current arrangements for shared office space as inappropriate for executing their marking and grading roles. Kuntz (2012) researched the faculty environment in relation to space, corroborating the findings here that academics needed to be afforded space which is conducive to completing ‘intellectual work’. It appears that faculty members require a more flexible use of space to facilitate separation when the task calls for a particular level of analytical activity. Yet such space needs to remain open to collegiate activities based on collaborative development of resources, research activity, and socialisation.

## 8.5 CONTRIBUTION TO CONTEMPORARY KNOWLEDGE

*"Perspective transformation is the process of becoming critically aware of how and why our presuppositions have come to constrain the way we perceive, understand, and feel about our world; of reformulating these assumptions to permit a more inclusive, discriminating, permeable and integrative perspective; and of making decisions or otherwise acting on these new understandings. More inclusive, discriminating permeable and integrative perspectives are superior perspectives that adults choose if they can because they are motivated to better understand the meaning of their and others' experience."*

(Mezirow, 1990:14 – my emphasis)

### Research Outcome 5

**RO5.** Messiness of marking and grading practice and troublesome knowledge.

Yorke (2011) explores assessment approaches through the dichotomies of the realist and relativist perspectives. These lie on a continuum as opposites with measurement as objective and as comparable to a specific standard at one end. At the other end of the spectrum lies a judgement about attainment being subjective and references against internal or individual standards which lecturers use for interpreting student performance. As Yorke (2011) highlights the first approach is seen as context free assessment whilst the second approach is viewed as context relevant. Criterion and normative or holistic referencing can also be viewed along this continuum. These nuances are in a way similar to the two schools of thought on phenomenology with Husserl and bracketing or pure situated at one end, and Heidegger and hermeneutics at the other (2.5.1, 2.5.2).

The real issue however lies with the utility, validity and reliability of the evaluation of the assessment artefact held by the lecturer. New or novice academics, potentially nearer to the lived experience of a student perhaps place more trust in a process, which as seen is fraught with pitfalls. Rather than being provided with a clear set of guidelines to follow, as is needed for newcomers (Hoffman 1998; Dreyfus 2004), they are given vague or unintelligible instructions for instance 'Jae' was told "...*have a go see what*

*you think ...*". Those new to the academy undertaking marking and grading potentially require more reassurance to execute their role.

It has been demonstrated in this research, by my interviewees, that the principles and practice of marking and grading are difficult to conceptualise and understand this, and could potentially be seen as being troublesome knowledge (Meyer and Land 2003). Novice evaluators remained confused about their role, how to execute it and also experienced an inability to enter into moderation discussions on an equal footing with those who appear to be more knowledgeable. Novice evaluators need to learn much about the particular ways in which their experienced colleagues think about and practice marking and grading. This does not appear to be as simple as being able to adopt a set of skills to be applied in a stepwise approach to facilitate performance of the act of marking and grading. Servage (2009) suggests that the intricate activities of teaching practice, and as is suggested here assessment and evaluation practice should be de-privatised and brought to the fore for utilisation amongst a community of practitioners. The suggestion presented from my research findings is to convene groups where those involved in evaluating a particular module engage in discussions pre and post evaluation. Participating in conversations about the practice of evaluation prior to commencement of the task will induct novices in a way which should facilitate revealing the 'know that' of marking and grading. Whilst a post evaluation discussion with the same group of lecturers will demonstrate much more of the 'know how' coming into view. Rowe and Martin (2014) in their study demonstrated that the use of polylogues amongst their research participants, set up as small and large discussion groups, prompted divergent interpretations to be revealed.

Managers may then more carefully select appropriate staff to mentor 'novices' specifically in making and grading, rather than assuming any lecturer leading a module can be effective at explaining this. Lecturers in my research with significant experience did not feel sure enough of their practice to support new academics as was identified by 'Harri' (CI). They may display the characteristics of expert practice, due to significant experience but has



not achieved the echelons of mindful competence (Figure 1). However it appears that experts find it almost impossible to communicate the 'know how' of marking and grading. The method by which lecturers develop the skills of judgement of assessment artefacts remains unclear. Data from my interviewees has cast marking and grading as troublesome knowledge. In light of this, professional development around this practice needs to concede that no single manual can be produced to explain marking and grading practice in a cohesive way. However issues around development of novices and newcomers to the local academy still remain, and should not be ignored.

#### **Research Outcome 6**

**RO6.** Interview - PA or CI as a way of examining marking or grading practice - could contribute to Continuing Professional Development, moderation and 2<sup>nd</sup> consideration to illuminate assessment practice and how evaluation decisions are reached.

Almost without exception interviewees participating in both data collection interview methods identified that discussing their marking and grading practice had been a cathartic and developmental experience. As identified at the outset of this thesis, marking and grading is an activity which is most often performed in isolation, as a separate event which is only discussed on a superficial level with colleagues. That discussion is generally the basis for moderation discussions, regarding what might be the correct mark and grade to award to an assessment artefact. There is limited, if any opportunity to explore in detail exactly what has been done and what has been considered to generate that mark or grade. Interviewees also identified that they had not experienced any opportunities to undertake professional development activities in which marking and grading was explored in detail. This includes how to use the marking and grading rubrics and any other tool which academics would be required to use for assessment. Interviewees completing a CI interview were encouraged to reflect in depth on a particular piece of work, recalling the assessment artefact, the particular task and elements within it that supported the decisions they made regarding it. This

close exploration encourages interviewees to think carefully about all aspects of their judgement as to the value of the artefact and what if anything impacted upon this. Talking in detail about their assessment decision in relation to one assessment artefact appeared to facilitate interviewees becoming more thoughtful on the process, and mindful of the implications of their judgement illustrating reflection on action (Schön 1995).

Whereas for the interviewees completing a PA interview, once they were clear about what they were being asked to talk out loud about, were able to talk of and about their evaluative decisions. Whilst the PA part was in progress interviewees vocalised thoughts and feelings about the artefact. Following this, on entering a post event discussion, interviewees reflected on their thinking and decision making. It was at this point that the PA interviewees explored what they recalled thinking about how they made those decisions.

#### **Research Outcome 7**

**RO7.** Experience recast as expertise in marking and grading; being able to perform the marking or grading task with limited conscious effort embodies unconscious competence. Evaluators can complete marking and grading without mindful attention.

The ability to 'know how' and to be able to explain the components of this 'know how' to others who have no or less knowledge is not shared by everyone. As alluded to earlier in this chapter marking and grading is a threshold concept cast as troublesome knowledge. It remains something difficult to understand and difficult to articulate even following passing through a gateway, being able to work with the 'know how' having integrated the 'know that'. An example of an individual with this level of expertise would be a driving instructor, who has the know how in being able to competently drive, but also is then able to explain all of the elements of this task effectively enough to transmit this knowledge to a novice or learner driver. Not everyone who can drive competently is able to use this knowledge to

support the development of this skill in others. Evidence can be identified, that this holds true in marking and grading contexts, from some interviewees in my research findings. There were some interviewees who found it difficult to explain how to complete the task to their more junior colleagues. With one interviewee specifically avoiding such conversations 'Harri' (CI)

*“... you don't have the time to ... go through the thing ... and of course if you don't feel overly confident about your own marking ...”*

It appears confidence in own abilities to complete the process goes some way towards facilitating evaluators in preparing new team members for their evaluative role. Learning to mark or grade is not simply a matter of undergoing formal and intentional education with appropriate social responses as superfluous, conveying processes through a use of artefacts and symbols. Processes of socialisation, in which novices are exposed to environments which facilitate adoption of appropriate attitudes and responses to the assessment artefacts, needs to occur. This is best achieved through provision of examples along with discussions that are inclusive of opportunities for articulating the marking and grading decisions made, inclusive of a rationale.

## 8.6 FROM NOVICE TO EXPERTISE: JOURNEY TO EXPERT PRACTICE

***“The simple things are also the most extraordinary things, and only the wise can see them.”***

Paulo Coelho

This research was facilitated by 26 lecturers in 4 HE institutions agreeing to participate in phenomenological data gathering via interviews. Employing two different methods enabled the lived experiences of lecturers in relation to marking and grading written assessment artefacts to be explored in depth, with both methods providing similar and distinct contributions to knowledge development. Findings reported above have facilitated taking a new perspective culminating in the development of a set of models for developing lecturer marking and grading practice, the ‘know that’ and ‘know how’. Elements in the model A (Figure 18) presented below are not new, but are in fact existing components of marking and grading practice. The claim to originality comes from how for the first time these elements have been surfaced and are being considered together as a potential way to support the development of new staff.

The close exploration of marking and grading practice has allowed for consideration of what constitutes expertise. Components of expert marking and grading practice identified through interactions with interviewees are portrayed in Model A (below).

### 8.6.1 DEVELOPING EXPERTISE IN MARKING AND GRADING PRACTICE:

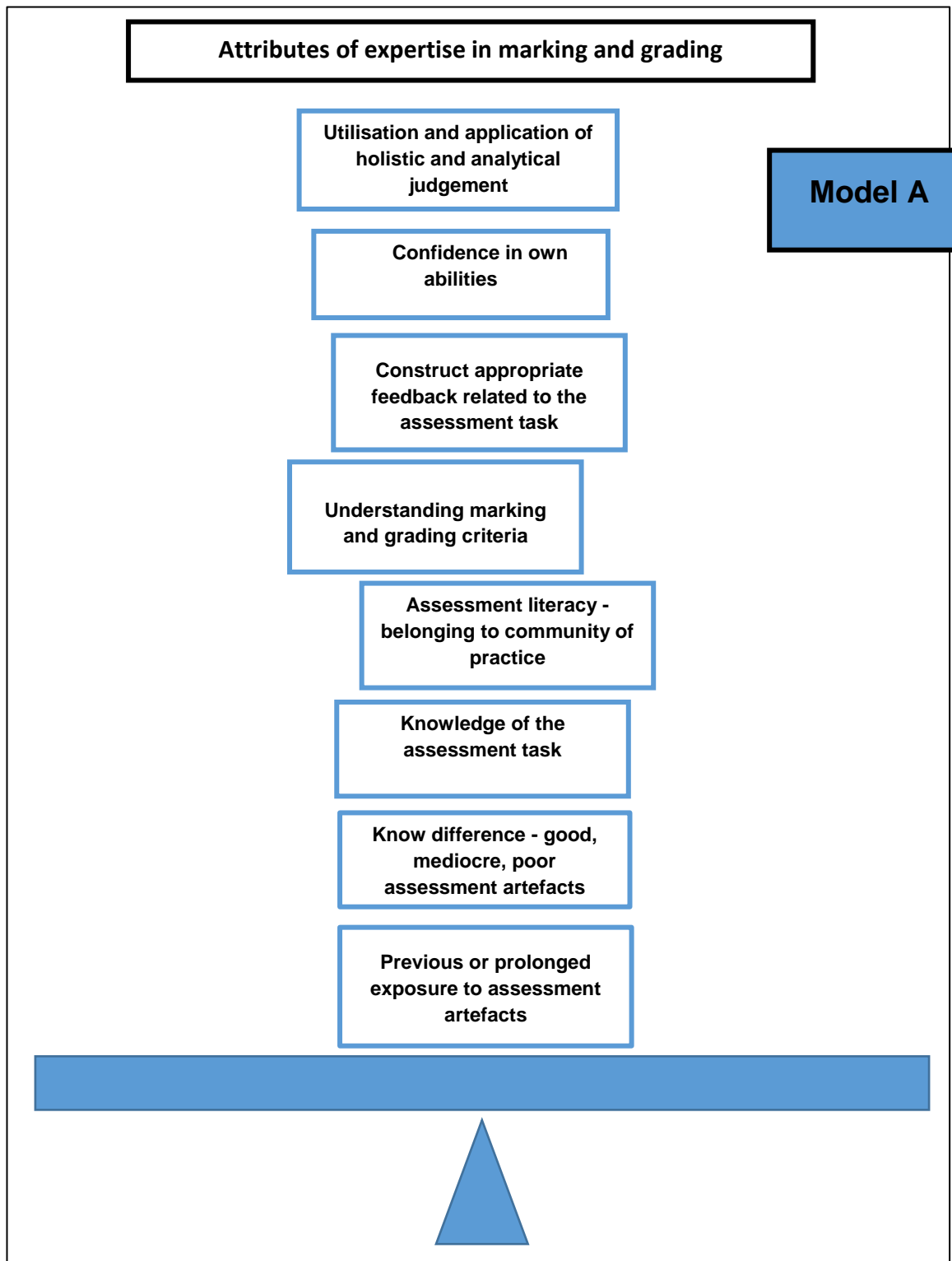
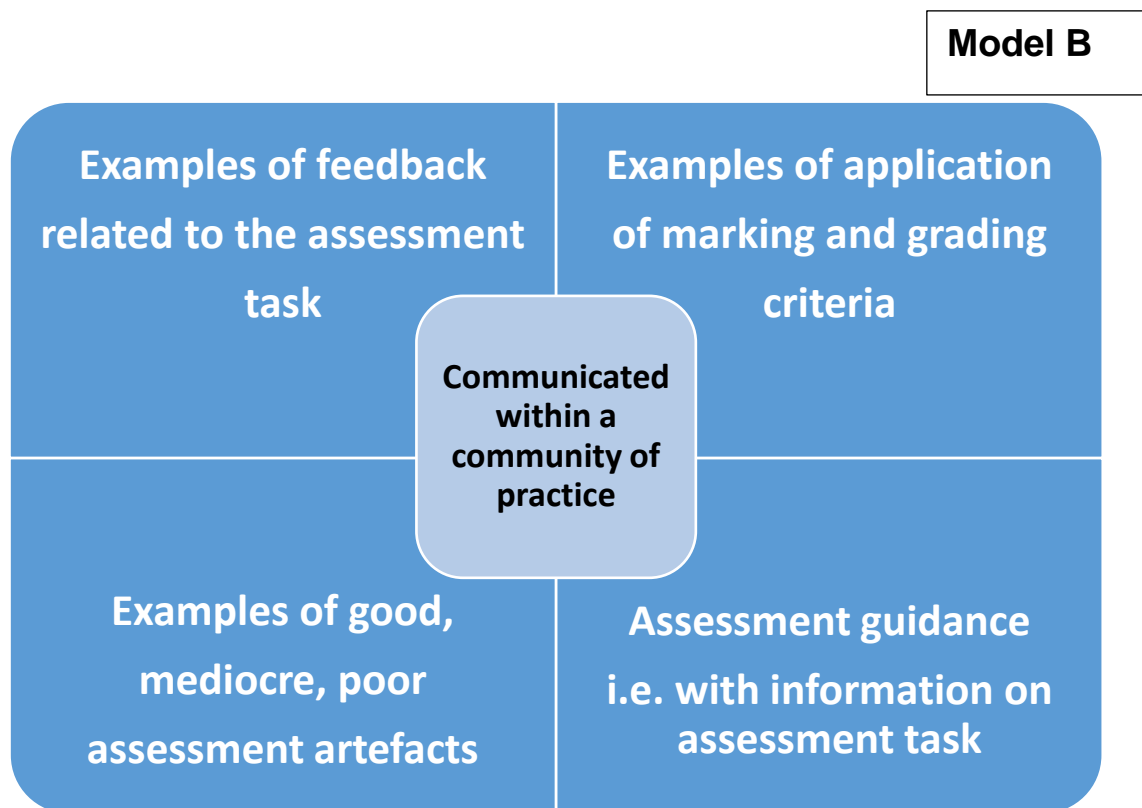


Figure 18 Attributes of expertise in marking and grading

Model A depicts all the evaluator attributes of a lecturer portraying expertise in marking and grading practice, is depicted as balanced on a pivot with changes suggested above interfering with this equilibrium. This model depicts a scale where one could easily move from being an 'expert' to becoming a novice when any change in the marking and grading environment occurs. These changes can be instigated by moving to a new department or new institution, losing membership of a familiar community of practice. Modifications which can occur through new programme and curriculum development incorporating new assessment approaches result in the loss of local assessment intelligence. Changes to marking and grading rubrics can also result in the loss of expertise in utilising the tool effectively, rendering the experienced evaluator to the position of advanced beginner, or novice.

Model B depicts the fundamental components required by a novice on introduction to marking and grading practice.



**Figure 19 Fundamental requirements for novice evaluators**

Model B proposes that in order to quickly develop competence in novice evaluators, they need to be immersed in the world of marking and grading practice. Using the four fundamentals in Model B will ensure that novice evaluators have a basic understanding of what they will need to do to mark and grade assessment artefacts from a particular module or assessment task. This model in no way suggests the development of expertise in marking and grading developing in a shorter time period than without their use. However the components of Model B propose the basic level of information required by novice academics, which is often missing from marking and grading induction as identified by my interviewees. Becoming steeped in a community of marking and grading practice could be accommodated by participation in pre evaluation and post evaluation conversations within a local team. This approach will support utilisation of the four fundamentals (Figure 19) and should include participation in moderation decisions which is an imperative step to learning about all elements of marking and grading practice. Using such an approach will ensure that novice evaluators are immersed in their local marking and grading community of practice.

Evidence of the need for more support for novice evaluators can be found within the excerpts from interviewees in this research. Whilst they are all no longer 'novice' evaluators, many clearly felt under prepared and unsupported for their evaluator role:

*Mel (CI) "... as a new member of staff ... there's your marking, off you go ..."*

*Ali (PA) "... I think when I started I found I really struggled with it ..."*

*Mel (CI) "... Will someone buddy up ... and go through it ... no one can ..."*

*Harri (CI) "...learnt their ways ... was right or not who knows ..."*

*Toni (PA) "... I feel that if no one else checks ..."*

These excerpts provide evidence of the need for support in the formative period whilst developing a raft of skills on entering the academy. Even though all the interviewees above have had a previous professional role, in

health care practice, which involved evaluating student performance and evaluating the outcome of care provided to patients, they still require support in developing a new raft of skills for academic practice.



## 8.7 SUMMARY

***“We delight in the beauty of the butterfly, but rarely admit the changes it has gone through to achieve that beauty.”***

Maya Angelou

The only thing that remains is to draw this thesis to a close. As is hinted by Maya Angelou (above), this chapter has afforded an opportunity to review the changes that have occurred, impacting upon my perspective on marking and grading written assessment artefacts by lecturers in higher education practice. The closing sections of this chapter focused on exploring the research outcomes identified at the end of Chapter 7 which had not previously been considered in the context of marking and grading by HE lecturers. This was structured into 3 distinct sections and began by discussing the main findings, the more concrete activity of marking and grading, in more general terms. It then went on to incorporate the first four research outcomes (RO1-4) related to the use and validity of marking and grading criteria and rubrics, the utility of establishing of CoP for marking and grading, the duality of health academics in marking and grading and their propensity for protecting practice standards, and working environments and its impact on practical aspects as well as considering work-life balance.

The second section focused on contributions to conceptual contemporary knowledge about marking and grading relevant to health academics. The final three research outcomes (RO5-7) commence with examining marking and grading as a threshold concept, acknowledging the difficulties in understanding and articulating the know-how of this activity. Then identifying a potential to utilise PA or CI for exploring marking and grading practice, either for teaching the know-that and the know-how to ‘novices’ and for continuing professional development for advanced beginners to experts as defined by Dreyfus (2004). The final research outcome identifies the utility of conceptualising expertise in marking and grading practice.

The third section of this concluding chapter focused on presenting my model identifying the features of expertise in marking and grading practice, drawing out four separate but fundamental components for developing 'novice' lecturers. The approach suggested must be nested within a local CoP and would provide support and opportunity for development, meeting the needs voiced by my interviewees. As highlighted by 'Mel' (CI): "...we have repeatedly said we'd like a workshop on marking. We'd like marking. We'd like marking. And no one would come ...". There is now potential for novice academics to be well prepared, furnished with skills and confidence in the processes for undertaking a key aspect of the role of an academic in HE.

## 8.8 RECOMMENDATIONS

***"The darkest hour is just before the dawn."***

Paulo Coelho

These have arisen from all that has gone before, whilst the list of recommendations are few, they could provide far reaching benefits well beyond the completion of any post-doctoral work.

- Investigate utility of the models for developing increased competence more quickly in marking and grading for novice evaluators (Model B).
- Developing marking and grading 'know that' as well as 'know how':
  - Accept and discuss marking and grading as a threshold concept, more likely to put novice lecturers at ease.
  - Develop shared understanding of marking and grading criteria or rubrics through discussion.
  - Instigate pre and post evaluation discussions with all lecturers involved in marking - develop a local CoP.
  - Discuss standards to be used for marking and grading to facilitate the use of overt rather than tacit personal and or professional criteria for evaluation - local community of practice.
- Use the models as a basis for exploring the marking and grading criteria or rubric to be applied, ensuring appropriate criteria are being

applied to evaluate assessment artefacts. Using inappropriate criteria or rubric could invalidate the final evaluation reducing its validity and interrater reliability.

- Develop a programme of continuing professional development which incorporates the use of PA or CI for closer examination of marking and grading practice.
- Development of new rubrics should be subject to scrutiny of a panel of academics, preferably not in the local CoP as the criteria or rubric should be self-explanatory. The use of an External Examiner could be seen as a way to provide objective scrutiny, or of inducting the individual into the local CoP.
- Consideration of the impact of shared office environments on the work of lecturers in relation to marking and grading.

# EPILOGUE

I have now reached the end of what has been an epic journey, one which was at times endured and, yet at others, I was able to conquer my fears. It seemed that those fears were many in the beginning, when I possessed limited knowledge of completing an empirical project, including understanding of research philosophy, methodology and methods. The greatest fear was probably one common amongst PhD students, would the final findings and conclusions result in the uncovering of something original? However my hopes for demystifying the marking and grading process have come to fruition. Above all I now have amassed concrete evidence demonstrating that marking and grading presents as troublesome knowledge for those new to academia. Therefore my contribution to contemporary knowledge on marking and grading is the development and evolution of a tool for the purpose of surfacing the components of marking and grading to support and develop educational practice. Using these models in the preparation and development of 'novice' and experienced academics for marking and grading practice, aims to reduce the anxiety experienced by this aspect of a lecturer's role. The next stage of my journey will be to champion its introduction in my own and other institutions with an intention to evaluate the outcomes.

***"I have no special talents. I am only passionately curious"***

Albert Einstein (n.d.)

# REFERENCE LIST

- Adie, L., Lloyd, M. and Beutel, D. (2013) Identifying discourses of moderation in higher education. *Assessment & Evaluation in Higher Education* 38 (8), 968-977.
- Aitken, L. M. and Mardegan, K. J. (2000) "Thinking Aloud": Data Collection in the Natural Setting. *Western Journal of Nursing Research* 22 (7), 841-853.
- Ajjawi, R. and Higgs, J. (2007) Using hermeneutic phenomenology to investigate how experienced practitioners learn to communicate clinical reasoning. *The Qualitative Report* 12 (4), 612-638.
- Ajzen, I. (2002) Perceived behavioral control, Self-Efficacy, locus of control, and the theory of planned Behavior<sup>1</sup>. *Journal of applied social psychology* 32 (4), 665-683.
- Allen, A. (2002) Power, Subjectivity, and Agency: Between Arendt and Foucault. *International Journal of Philosophical Studies* 10 (2), 131-149.
- Anderson, G. L. and Herr, K. (1999) The new paradigm wars: Is there room for rigorous practitioner knowledge in schools and universities? *Educational researcher* 28 (5), 12-40.
- Andrade, H. G. (2005) Teaching with rubrics: The good, the bad, and the ugly. *College teaching* 53 (1), 27-31.
- Annala, J. and Mäkinen, M. (2016) Communities of practice in higher education: contradictory narratives of a university-wide curriculum reform. *Studies in Higher Education*, 1-17.
- Annells, M. (1996) Hermeneutic phenomenology: philosophical perspectives and current use in nursing research. *Journal of Advanced Nursing* 23 (4), 705-713.
- Armitage, G., Cracknell, A., Forrest, K. and Sandars, J. (2011) Twelve tips for implementing a patient safety curriculum in an undergraduate programme in medicine. *Medical teacher* 33 (7), 535-540.
- Ashworth, M., Bloxham, S. and Pearce, L. (2010) Examining the tension between academic standards and inclusion: the impact on marking of individuals' frameworks for assessment. *Studies in Higher Education* 35 (2).
- Atkinson, R. C. and Shiffrin, R. M. (1968) Human memory: A proposed system and its control processes. *The psychology of learning and motivation* 2, 89-195.

- Azzur (2010) *Mindful Leader*. <http://www.azzur.co.uk/blog/?p=266>  
Accessed 30/3/14.
- Baddeley, A. D. (1997) *Human memory: Theory and practice*. Psychology Press.
- Baehr, J. S. (2006) *A Priori and A Posteriori*.  
<http://www.iep.utm.edu/apriori/>; <http://www.iep.utm.edu/> Accessed 23/4/2015.
- Ball, E., Franks, H., Jenkins, J., McGrath, M. and Leigh, J. (2009) Annotation is a valuable tool to enhance learning and assessment in student essays. *Nurse education today* 29 (3), 284-291.
- Banbury, S. P. and Berry, D. C. (2005) Office noise and employee concentration: Identifying causes of disruption and potential improvements. *Ergonomics* 48 (1), 25-37.
- Baume, D., Yorke \*, M. and Coffey, M. (2004) What is happening when we assess, and how can we use our understanding of this to improve assessment? *Assessment & Evaluation in Higher Education* 29 (4), 451-477.
- Bazeley, P. (2009) Analysing qualitative data: More than 'identifying themes'. *Malaysian Journal of Qualitative Research* 2 (2), 6-22.
- Bazeley, P. (2013) *Qualitative Data Analysis: Practical Strategies*. London: Sage Publications.
- Bazeley, P. and Jackson, K. (2013) *Qualitative data analysis with NVivo*. Sage Publications Limited.
- Beatty, P. C. and Willis, G. B. (2007) Research Synthesis: The Practice of Cognitive Interviewing. *Public Opinion Quarterly* 71 (2), 287-311.
- Benbunan-Fich, R. (2001) Using protocol analysis to evaluate the usability of a commercial web site. *Information & Management* 39 (2), 151-163.
- Benner, P. (1984) From novice to expert. *Menlo Park*.
- Berliner, D. C. (2004) Describing the Behavior and Documenting the Accomplishments of Expert Teachers. *Bulletin of Science, Technology & Society* 24 (3), 200-212.
- Bettany-Saltikov, J., Kilinc, S. and Stow, K. (2009) Bones, boys, bombs and booze: an exploratory study of the reliability of marking dissertations across disciplines. *Assessment & Evaluation in Higher Education* 34 (6), 621-639.

- Beyer, C. (2013) "Edmund Husserl", *The Stanford Encyclopedia of Philosophy, Summer2013 Edition*.  
<<http://plato.stanford.edu/archives/sum2013/entries/husserl/>>  
Accessed 18/5/13.
- Biggs, J. (1996) Enhancing teaching through constructive alignment. *Higher Education* 32 (3), 347-364.
- Biggs, J. and Tang, C. (2009) Applying constructive alignment to outcomes-based teaching and learning. 2010.  
[https://intranet.tudelft.nl/fileadmin/Files/medewerkersportal/TBM/Onde\\_rwijsdag\\_2014/What-is-ConstructiveAlignment.pdf](https://intranet.tudelft.nl/fileadmin/Files/medewerkersportal/TBM/Onde_rwijsdag_2014/What-is-ConstructiveAlignment.pdf).
- Black, P. and Wiliam, D. (2009) Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability(formerly: Journal of Personnel Evaluation in Education)* 21 (1), 5.
- Blackaby, D. and Frank, J. (2000) Ethnic and Other Minority Representation in UK Academic Economics. *The Economic Journal* 110 (464), 293-311.
- Bloxham, S. (2009) Marking and moderation in the UK: false assumptions and wasted resources. *Assessment & Evaluation in Higher Education* 34 (2), 209-220.
- Bloxham, S. and Boyd, P. (2012) Accountability in grading student work: securing academic standards in a twenty-first century quality assurance context. *British Educational Research Journal* 38 (4), 615-634.
- Bloxham, S., Boyd, P. and Orr, S. (2011) Mark my words: the role of assessment criteria in UK higher education grading practices. *Studies in Higher Education* 36 (6), 655-670.
- Bloxham, S., den-Outer, B., Hudson, J. and Price, M. (2015) Let's stop the pretence of consistent marking: exploring the multiple limitations of assessment criteria. *Assessment & Evaluation in Higher Education*, 1-16.
- Bloxham, S., Hughes, C. and Adie, L. (2016) What's the point of moderation? A discussion of the purposes achieved through contemporary moderation practices. *Assessment & Evaluation in Higher Education* 41 (4), 638-653.
- Bloxham, S. and West, A. (2004) Understanding the rules of the game: marking peer assessment as a medium for developing students' conceptions of assessment. *Assessment & Evaluation in Higher Education* 29 (6), 721-733.

- Bloxham, S. and West, A. (2007) Learning to write in higher education: students perceptions of an intervention in developing understanding of assessment criteria. *Teaching in Higher Education* 12 (1), 77 - 89.
- Boell, S. K. and Cecez-Kecmanovic, D. (2014) A Hermeneutic Approach for Conducting Literature Reviews and Literature Searches. *Communications of the Association for Information Systems* 34, 257-286.
- Boud, D. and Falchikov, N. (2006) Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education* 31 (4), 399-413.
- Boyd, P. (2010) Academic induction for professional educators: supporting the workplace learning of newly appointed lecturers in teacher and nurse education. *International Journal for Academic Development* 15 (2), 155-165.
- Boyd, P., Ashworth, M., Bloxham, S. and Orr, S. (2009) Grading Student Work: Using think aloud to investigate the assessment practices of university lecturers. In *British Educational Research Association*. Manchester. University of Manchester. [http://www.bera.ac.uk/bera2009/docs/BERA2009\\_0560.pdf](http://www.bera.ac.uk/bera2009/docs/BERA2009_0560.pdf)
- Boyd, P. and Bloxham, S. (2013) A situative metaphor for teacher learning: the case of university tutors learning to grade student coursework. *British Educational Research Journal*.
- Boyd, P. and Bloxham, S. (2014) A situative metaphor for teacher learning: the case of university tutors learning to grade student coursework. *British Educational Research Journal* 40 (2), 337-352.
- Brennan, D. J. (2008) University student anonymity in the summative assessment of written work. *Higher Education Research & Development* 27 (1), 43 - 54.
- Bridges, D. (2001) The Ethics of Outsider Research. *Journal of Philosophy of Education* 35 (3), 371-386.
- Bridges, D. (2009) Education and the possibility of outsider understanding. *Ethics and Education* 4 (2), 105-123.
- Bridges, P., Bourdillon, B., Collymore, D., Cooper, A., Fox, W., Haines, C., Turner, D., Woolf, H. and Yorke, M. (1999) Discipline-related Marking Behaviour Using Percentages: a potential cause of inequity in assessment. *Assessment & Evaluation in Higher Education* 24 (3), 285-300.



- Bromley, M. (2011) *Clinical Human Factors Group Working with clinical professionals and managers to make healthcare safer*.  
[www.http://chfg.org/](http://chfg.org/) Accessed 24 November 2015.
- Brooker, R. and Macpherson, I. (1999) Communicating the processes and outcomes of practitioner research: an opportunity for self-indulgence or a serious professional responsibility? *Educational Action Research* 7 (2), 207-221.
- Brooks, V. (2004) Double marking revisited. *British Journal of Educational Studies* 52 (1), 29-46.
- Brooks, V. (2012) Marking as judgment. *Research Papers in Education* 27 (1), 63-80.
- Burgess, T. F. (1996) Planning the academic's workload: Different approaches to allocating work to university academics. *Higher education* 32 (1), 63-75.
- Burgess, T. F., Lewis, H. A. and Mobbs, T. (2003) Academic Workload Planning Revisited. *Higher Education* 46 (2), 215-233.
- Burr, V. (2003) *Social constructionism*. Psychology Press.
- Butler Shay, S. (2004) The assessment of complex performance: A socially situated interpretive act. *Harvard Educational Review* 74 (3), 307-329.
- Carless, D. (2006) Differing perceptions in the feedback process. *Studies in Higher Education* 31 (2), 219 - 233.
- Carpentier, V. (2008) Quantitative sources for the history of education. *History of Education* 37 (5), 701-720.
- Carthey, J. and Clarke, J. (2010) Implementing Human Factors in Healthcare.  
[http://www.improvementacademy.org/documents/Projects/human\\_factors/Human-Factors-How-to-Guide-v1.2.pdf](http://www.improvementacademy.org/documents/Projects/human_factors/Human-Factors-How-to-Guide-v1.2.pdf)
- Chenail, R. J. (2011) Interviewing the investigator: Strategies for addressing instrumentation and researcher bias concerns in qualitative research. *The Qualitative Report* 16 (1), 255-262.
- Clouder, L. (2005) Caring as a 'threshold concept': transforming students in higher education into health(care) professionals. *Teaching in Higher Education* 10 (4), 505-517.
- Constanti, P. and Gibbs, P. (2004) Higher education teachers and emotional labour. *International Journal of Educational Management* 18 (4), 243-249.

- Coupal, L. (2005) Practitioner-Research and the Regulation of Research Ethics: The Challenge of Individual, Organizational, and Social Interests. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research; Vol 6, No 1 (2005): Secondary Analysis of Qualitative Data.*
- Craik, F. I. M. and Lockhart, R. S. (1972) Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior* 11 (6), 671-684.
- Creswell, J. W. (2012) *Qualitative inquiry and research design: Choosing among five approaches.* SAGE Publications, Incorporated.
- Crisp, B. R. (2007) Is it worth the effort? How feedback influences students subsequent submission of assessable work. *Assessment & Evaluation in Higher Education* 32 (5), 571 - 581.
- Crisp, V. (2008) The validity of using verbal protocol analysis to investigate the processes involved in examination marking. *Research in Education* (79), 1-12.
- Crisp, V. (2010a) Judging the Grade: Exploring the Judgement Processes Involved in Examination Grading Decisions. *Evaluation & Research in Education* 23 (1), 19-35.
- Crisp, V. (2010b) Towards a model of the judgement processes involved in examination marking. *Oxford Review of Education* 36 (1), 1-21.
- Crisp, V. and Johnson, M. (2007) The use of annotations in examination marking: opening a window into markers' minds. *British Educational Research Journal* 33 (6), 943-961.
- Crisp, V. and Novaković, N. (2009) Are all assessments equal? The comparability of demands of college-based assessments in a vocationally related qualification. *Research in Post-Compulsory Education* 14 (1), 1-18.
- Cromby, J. and Nightingale, D. (1999) *Social Constructionist Psychology: a critical analysis of theory and practice.* Buckingham: Open University Press.
- Crotty, M. (1998) *The Foundations of Social Research. Meaning and perspective in research process.* Australia: Sage.
- Davies, M. B. and Hughes, N. (2014) *Doing a successful research project: Using qualitative or quantitative methods.* Palgrave Macmillan.
- Dawson, P. (2017) Assessment rubrics: towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education* 42 (3), 347-360.

de Witt, L. and Ploeg, J. (2006) Critical appraisal of rigour in interpretive phenomenological nursing research. *Journal of Advanced Nursing* 55 (2), 215-229.

Dearing, S. R. and NCIHE (1997) *Higher Education in the Learning Society: Reports of the National Committee*. NCIHE.

Dowling, M. (2007) From Husserl to van Manen. A review of different phenomenological approaches. *International Journal of Nursing Studies* 44 (1), 131-142.

Drennan, J. (2003) Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *Journal of advanced nursing* 42 (1), 57-63.

Dreyfus, S. E. (2004) The five-stage model of adult skill acquisition. *Bulletin of science, technology & society* 24 (3), 177-181.

Dreyfus, S. E. and Dreyfus, H. L. (1980) *A five-stage model of the mental activities involved in directed skill acquisition*. DTIC Document.

Dubey, P. and Geanakoplos, J. (2010) Grading exams: or ? *Games and Economic Behavior* 69 (1), 72-94.

Duguid, P. (2005) "The Art of Knowing": Social and Tacit Dimensions of Knowledge and the Limits of the Community of Practice. *The Information Society* 21 (2), 109-118.

Dunn, L., Morgan, C. and Parry, S. (2002) Seeking quality in criterion referenced assessment. In *EARLi Special Interest Group on Assessment and Evaluation*. Northumbria. Education-line. <http://www.leeds.ac.uk/educol/documents/00002257.htm>

Ecclestone, K. and Swann, J. (1999) Litigation and Learning: Tensions in improving university lecturers' assessment practice. *Assessment in Education: Principles, Policy & Practice* 6 (3), 377-389.

Elander, J. and Hardman, D. (2002) An application of judgment analysis to examination marking in psychology. *British Journal of Psychology* 93 (3), 303-328.

Entwistle, N. (2008) Threshold concepts and transformative ways of thinking within research into higher education. *Threshold concepts within the disciplines*, 21-35.

Equality Act (SI 6)

<http://www.legislation.gov.uk/ukpga/2010/15/part/6/chapter/2>

Ericsson, K. A. (2004) Deliberate Practice and the Acquisition and Maintenance of Expert Performance in Medicine and Related Domains. *Academic Medicine* 79 (10).

- Ericsson, K. A. (2006) Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative tasks. *The Cambridge handbook of expertise and expert performance*, 223-241.
- Ericsson, K. A. and Simon, H. A. (1980) Verbal reports as data. *Psychological review* 87 (3), 215.
- Ericsson, K. A. and Simon, H. A. (1993) *Protocol analysis: verbal reports as data*. Revised edition. MIT press.
- Finlay, L. (2002) "Outing" the researcher: The provenance, process, and practice of reflexivity. *Qualitative Health Research* 12 (4), 531-545.
- Finlay, L. (2006) 'Rigour', 'ethical integrity' or 'artistry'? Reflexively reviewing criteria for evaluating qualitative research. *The British Journal of Occupational Therapy* 69 (7), 319-326.
- Finlay, L. and Gough, B. (2008) *Reflexivity: A practical guide for researchers in health and social sciences*. John Wiley & Sons.
- Fisher, R. P. (1995) Interviewing victims and witnesses of crime. *Psychology, Public Policy, and Law* 1 (4), 732-764.
- Fisher, R. P., Geiselman, R. E. and Amador, M. (1989) Field test of the cognitive interview: Enhancing the recollection of the actual victims and witnesses of crime. *Journal of Applied Psychology* 74 (5), 722-727.
- Fleming, V., Gaidys, U. and Robb, Y. (2003) Hermeneutic research in nursing: developing a Gadamerian-based research method. *Nursing Inquiry* 10 (2), 113-120.
- Flynn, B. (2011) *Maurice Merleau-Ponty*.
- Fox, M., Martin, P. and Green, G. (2007) *Doing practitioner research*. Sage.
- Fuller, A., Hodkinson, H., Hodkinson, P. and Unwin, L. (2005) Learning as peripheral participation in communities of practice: a reassessment of key concepts in workplace learning. *British Educational Research Journal* 31 (1), 49-68.
- Fuller\*, M., Bradley, A. and Healey, M. (2004) Incorporating disabled students within an inclusive higher education environment. *Disability & Society* 19 (5), 455-468.
- Funkesson, K. H., Anbäcken, E.-M. and Ek, A.-C. (2007) Nurses' reasoning process during care planning taking pressure ulcer prevention as an example. A think-aloud study. *International Journal of Nursing Studies* 44 (7), 1109-1119.

- Füredi, F. (2009) *Wasted: why education isn't educating*. Continuum International Publishing Group.
- Garcia, S. B. and Guerra, P. L. (2004) Deconstructing deficit thinking: Working with educators to create more equitable learning environments. *Education and urban society* 36 (2), 150-168.
- Gardner, L. (2012) From Novice to Expert: Benner's legacy for nurse education. *Nurse Education Today* 32 (4), 339-340.
- Garrow, A. and Tawse, S. (2009) An exploration of the assessment experiences of new academics as they engage with a community of practice in higher education. *Nurse Education Today* 29 (6), 580-584.
- Gascoigne, N. and Thornton, T. (2014) *Tacit knowledge*. Routledge.
- Geiselman, R. E., Fisher, R. P., MacKinnon, D. P. and Holland, H. L. (1986) Enhancement of Eyewitness Memory with the Cognitive Interview. *The American Journal of Psychology* 99 (3), 385-401.
- Gherardi, S. (2009) Introduction: The Critical Power of the 'Practice Lens'. *Management Learning* 40 (2), 115-128.
- Gibbs, G. and Simpson, C. (2004-05) Conditions Under Which Assessment Supports Students' Learning. *Learning & Teaching in Higher Education* (1), 28.
- Gillborn, D. (2010) The colour of numbers: surveys, statistics and deficit-thinking about race and class. *Journal of Education Policy* 25 (2), 253-276.
- Gillett, A. and Hammond, A. (2009) Mapping the maze of assessment: An investigation into practice. *Active Learning in Higher Education* 10 (2), 120-137.
- Gillon, R. (1994) Medical ethics: four principles plus attention to scope. *BMJ: British Medical Journal* 309 (6948), 184.
- Goldman, A. I. (2016) Expertise. *Topoi*, 1-8.
- Gourlay, L. (2011) 'I'd landed on the moon': a new lecturer leaves the academy. *Teaching in Higher Education* 16 (5), 591-601.
- Grainger, P., Purnell, K. and Zipf, R. (2008) Judging quality through substantive conversations between markers. *Assessment & Evaluation in Higher Education* 33 (2), 133-142.
- Greasley, P. and Cassidy, A. (2009) When it comes round to marking assignments: how to impress and how to 'distress' lecturers .... *Assessment & Evaluation in Higher Education* 35 (2), 173-189.

- Hack, C. (2015) Analytical rubrics in higher education: A repository of empirical data. *British Journal of Educational Technology* 46 (5), 924-927.
- Halcomb, E. J. and Davidson, P. M. (2006) Is verbatim transcription of interview data always necessary? *Applied Nursing Research* 19 (1), 38-42.
- Hand, L. and Clewes, D. (2000) Marking the Difference: An investigation of the criteria used for assessing undergraduate dissertations in a business school. *Assessment & Evaluation in Higher Education* 25 (1), 5-21.
- Handley, K., den Outer, B. and Price, M. (2013) Learning to mark: exemplars, dialogue and participation in assessment communities. *Higher Education Research & Development* 32 (6), 888-900.
- Hannan, A. and Silver, H. (2004) ENQUIRY INTO THE NATURE OF EXTERNAL EXAMINING Final report. *Higher Education Academy*. Retrieved June 23, 2006.
- Hartley, J., Trueman, M., Betts, L. and Brodie, L. (2006) What price presentation? The effects of typographic variables on essay grades. *Assessment & Evaluation in Higher Education* 31 (5), 523-534.
- Herbert, I. P., Joyce, J. and Hassall, T. (2014) Assessment in higher education: The potential for a community of practice to improve inter-marker reliability. *Accounting Education* 23 (6), 542-561.
- Hewitt, J. (2007) Ethical Components of Researcher—Researched Relationships in Qualitative Interviewing. *Qualitative health research* 17 (8), 1149-1159.
- Higgins, R., Hartley, P. and Skelton, A. (2001) Getting the Message Across: the problem of communicating assessment feedback. *Teaching in Higher Education* 6 (2), 269 - 274.
- Higgins, R., Hartley, P. and Skelton, A. (2002) The Conscientious Consumer: reconsidering the role of assessment feedback in student learning. *Studies in Higher Education* 27 (1), 53 - 64.
- Hodgkinson, G. P., Langan-Fox, J. and Sadler-Smith, E. (2008) Intuition: A fundamental bridging construct in the behavioural sciences. *British Journal of Psychology* 99 (1), 1-27.
- Hoffman, R. R. (1998) How can expertise be defined? Implications of research from cognitive psychology. *Exploring expertise*, 81-100.
- Holstein, J. A. and Gubrium, J. F. (1995) *The active interview*. Vol. 37. London: Sage.

- Hornby, W. I. N. (2003) Assessing Using Grade-related Criteria: A single currency for universities? *Assessment & Evaluation in Higher Education* 28 (4), 435-454.
- House of Commons. (2009) *Widening participation in higher education: Fourth Report of Session 2008-2009*. Committee, P. A. London: The Stationery Office Limited.  
<http://www.publications.parliament.uk/pa/cm200809/cmselect/cmpubacc/226/9780215526557.pdf>
- Hudson, J., Bloxham, S., den Outer, B. and Price, M. (2015) Conceptual acrobatics: talking about assessment standards in the transparency era. *Studies in Higher Education*, 1-15.
- Hunter, B. and Smith, P. (2007) Emotional labour: Just another buzz word? *International Journal of Nursing Studies* 44 (6), 859-861.
- Hunter, K. and Docherty, P. (2009) Reducing variation in the assessment of student writing. *Assessment & Evaluation in Higher Education* 36 (1), 109-124.
- Hycner, R. H. (1985) Some guidelines for the phenomenological analysis of interview data. *Human studies* 8 (3), 279-303.
- International, G. T. (2012) *Conscious Competence Model*. 2016. Solana Beach, California: GT International.  
<http://www.gordontraining.com/free-workplace-articles/learning-a-new-skill-is-easier-said-than-done/#>
- Jacelon, C. S. and O'Dell, K. K. (2005) Analyzing qualitative data. *Urologic Nursing* 25 (3), 217-220.
- Jawitz, J. (2009) Learning in the academic workplace: the harmonization of the collective and the individual habitus. *Studies in Higher Education* 34 (6), 601-614.
- Jobs.ac.uk (2013) *Higher Education Job Seeking Careers Advice*.  
<http://www.jobs.ac.uk/careers-advice/jobseeking-tips/489/top-10-tips-from-hr-professionals/> Accessed 9/12/13.
- Jonsson, A. and Svingby, G. (2007) The use of scoring rubrics: Reliability, validity and educational consequences. *Educational research review* 2 (2), 130-144.
- Kinman, G. and Jones, F. (2008) A life beyond work? Job demands, work-life balance, and wellbeing in UK academics. *Journal of Human Behavior in the Social Environment* 17 (1-2), 41-60.
- Knight, P. (1995) *Assessment for learning in higher education*. Psychology Press.

- Koch, T. (1995) Interpretive approaches in nursing research: the influence of Husserl and Heidegger. *Journal of Advanced Nursing* 21 (5), 827-836.
- Koch, T. and Harrington, A. (1998) Reconceptualizing rigour: the case for reflexivity. *Journal of Advanced Nursing* 28 (4), 882-890.
- Kuntz, A. M. (2012) Reconsidering the workplace: faculty perceptions of their work and working environments. *Studies in Higher Education* 37 (7), 769-782.
- Kuusela, H. and Paul, P. (2000) A comparison of concurrent and retrospective verbal protocol analysis. *The American journal of psychology* (113), 387-404.
- Kvale, S. (2008) *InterViews: Learning The Craft Of Qualitative Research Interviewing* Author: Steinar Kvale, Svend Brinkmann, Publisher: Sag.
- Larrington, C. and Roger, L. (2002) Equal among first: Reviewing quality in Psychology. *Psychology Learning and Teaching* 2 (1), 6.
- Lave, J. and Wenger, E. (1991) *Situated learning: Legitimate peripheral participation*. Cambridge university press.
- Lave, J. and Wenger, E. (1998) Communities of practice. *Infed* 9, 8.  
<http://valenciacollege.edu/faculty/development/tla/documents/CommunityofPractice.pdf>
- Lawson, K. and Yorke, J. (2009) The development of moderation across the institution: a comparison of two approaches. *ATN Assessment Conference*. 2009. Citeseer.  
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.475.8063&rep=rep1&type=pdf#page=236>.
- Lea, M. R. and Street, B. V. (1998) Student writing in higher education: An academic literacies approach. *Studies in higher education* 23 (2), 157-172.
- Leathwood, C. (2005) Assessment policy and practice in higher education: Purpose, standards and equity. *Assessment & Evaluation in Higher Education* 30 (3), 307-324.
- An Act to amend the law relating to education 1988, c40* (c. 2) London, England:  
[http://www.legislation.gov.uk/ukpga/1988/40/part/IIhttp://www.legislation.gov.uk/ukpga/1988/40/pdfs/ukpga\\_19880040\\_en.pdf](http://www.legislation.gov.uk/ukpga/1988/40/part/IIhttp://www.legislation.gov.uk/ukpga/1988/40/pdfs/ukpga_19880040_en.pdf)
- Lillis, T. and Turner, J. (2001) Student Writing in Higher Education: Contemporary confusion, traditional concerns. *Teaching in Higher Education* 6 (1), 57-68.



- Lizzio, A. and Wilson, K. (2008) Feedback on assessment: students' perceptions of quality and effectiveness. *Assessment & Evaluation in Higher Education* 33 (3), 263-275.
- Lok, B., McNaught, C. and Young, K. (2016) Criterion-referenced and norm-referenced assessments: compatibility and complementarity. *Assessment & Evaluation in Higher Education* 41 (3), 450-465.
- Mackey, S. (2005) Phenomenological nursing research: methodological insights derived from Heidegger's interpretive phenomenology. *International Journal of Nursing Studies* 42 (2), 179-186.
- MacLaren, J., Haycock-Stuart, E., McLachlan, A. and James, C. (2016) Understanding pre-registration nursing fitness to practise processes. *Nurse Education Today* 36, 412-418.
- Madriaga, M., Hanson, K., Heaton, C., Kay, H., Newitt, S. and Walker, A. (2010) Confronting similar challenges? Disabled and non-disabled students' learning and assessment experiences. *Studies in Higher Education* 35 (6), 647-658.
- Malcolm, J. and Zukas, M. (2009) Making a mess of academic work: experience, purpose and identity. *Teaching in Higher Education* 14 (5), 495-506.
- Malouff, J. M., Emmerton, A. J. and Schutte, N. S. (2013) The risk of a halo bias as a reason to keep students anonymous during grading. *Teaching of Psychology*, 0098628313487425.
- Mason, J. (2002a) *Qualitative Researching*. SAGE Publications.
- Mason, J. (2002b) *Qualitative researching*. 2nd edition. London: Sage.
- McArthur-Rouse, F. J. (2008) From expert to novice: An exploration of the experiences of new academic staff to a department of adult nursing studies. *Nurse Education Today* 28 (4), 401-408.
- McConlogue, T. (2011) But is it fair? Developing students' understanding of grading complex written work through peer assessment. *Assessment & Evaluation in Higher Education* 37 (1), 113-123.
- McDermid, F., Peters, K., Daly, J. and Jackson, D. (2016) Developing resilience: Stories from novice nurse academics. *Nurse Education Today* 38, 29-35.
- McGregor, C. (2007) Self-fashioning through memoir: becoming an adult educator. *Teacher Development* 11 (1), 77-97.
- Meiklejohn, J., Phillips, C., Freedman, M. L., Griffin, M. L., Biegel, G., Roach, A., Frank, J., Burke, C., Pinger, L. and Soloway, G. (2012)

- Integrating mindfulness training into K-12 education: Fostering the resilience of teachers and students. *Mindfulness* 3 (4), 291-307.
- Memon, A. and Bull, R. (1991) The cognitive interview: Its origins, empirical support, evaluation and practical implications. *Journal of Community & Applied Social Psychology* 1 (4), 291-307.
- Memon, A. and Higham, P. A. (1999) A review of the cognitive interview. *Psychology, Crime and Law* 5 (1-2), 177-196.
- Meyer, J. and Land, R. (2003) *Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practising within the disciplines*. University of Edinburgh Edinburgh.
- Meyer, J. H. F. and Land, R. (editors) (2006) *Threshold concepts and troublesome knowledge*. Routledge Abingdon,, UK.
- Milligan, F. (1996) The use of criteria-based grading profiles in formative and summative assessment. *Nurse Education Today* 16 (6), 413-418.
- Monaghan, T. (2015) A critical analysis of the literature and theoretical perspectives on theory–practice gap amongst newly qualified nurses within the United Kingdom. *Nurse Education Today* 35 (8), e1-e7.
- Morgan, G. and Smircich, L. (1980) The Case for Qualitative Research. *The Academy of Management Review* 5 (4), 491-500.
- Moustakas, C. (1994) *Phenomenological research methods*. SAGE Publications, Incorporated.
- NatCen (2013) *Cognitive Testing*. London: <http://www.natcen.ac.uk/our-expertise/methods-expertise/cognitive-testing/national-institute-of-health-research/> Accessed 20/12/13.
- National Union of Students, N. (2008) *Mark My Words Not My Name*.
- Nolen, A. L. and Putten, J. V. (2007) Action Research in Education: Addressing Gaps in Ethical Principles and Practices. *Educational Researcher* 36 (7), 401-407.
- Northedge, A. (2003) Enabling Participation in Academic Discourse. *Teaching in Higher Education* 8 (2), 169 - 180.
- Nursing and Midwifery Council (2009) *Standards for pre registration midwifery education*. London: NMC, Nursing and Midwifery Council. [http://www.nmc-uk.org/Documents/NMC-Publications/nmcStandardsforPre\\_RegistrationMidwiferyEducation.pdf](http://www.nmc-uk.org/Documents/NMC-Publications/nmcStandardsforPre_RegistrationMidwiferyEducation.pdf)
- Nursing and Midwifery Council and NMC (2008) *The Code: Standards of conduct, performance and ethics for nurses and midwives*. London:

<http://www.nmc-uk.org/Documents/Standards/The-code-A4-20100406.pdf>

- Nursing and Midwifery Council and NMC (2016a) *Health and character guidance for AEs*. <https://www.nmc.org.uk/education/what-we-expect-of-educational-institutions/good-health-and-good-character-for-aeis/>: NMC.
- Nursing and Midwifery Council and NMC (2016b) *What is fitness to practice?* 1/3/2016. <https://www.nmc.org.uk/concerns-nurses-midwives/what-we-do/what-is-fitness-to-practise/>: NMC. <https://www.nmc.org.uk/concerns-nurses-midwives/what-we-do/what-is-fitness-to-practise/> Accessed 26/6/16.
- Nursing and Midwifery Council (NMC) (2015) *The code: professional standards of practice and behaviour for nurses and midwives*. NMC, London.
- Nursing and Midwifery Council. (2010) *Standards for Pre-Registration Nurse Education*. London, NMC.
- O'Donovan, B., Price, M. and Rust, C. (2001) The student experience of criterion-referenced assessment (through the introduction of a common criteria assessment grid). *Innovations in Education and Teaching International* 38 (1), 74-85.
- O'Donovan, B., Price, M. and Rust, C. (2004) Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education* 9 (3), 325-335.
- O'Donovan, B., Price, M. and Rust, C. (2008) Developing student understanding of assessment standards: a nested hierarchy of approaches. *Teaching in Higher Education* 13 (2), 205-217.
- Office for National Statistics, O. (2016) *Annual Survey of Hours and Earnings: 2015 Provisional Results*. <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/bulletins/annualsurveyofhoursandearnings/2015provisionalresults#earnings-by-occupation> Accessed 22/10/2016.
- Orb, A., Eisenhauer, L. and Wynaden, D. (2001) Ethics in Qualitative Research. *Journal of Nursing Scholarship* 33 (1), 93-96.
- Orr, K. and Bennett, M. (2009) Reflexivity in the co-production of academic-practitioner research. *Qualitative Research in Organizations and Management* 4 (1), 85-102.
- Orr, S. (2007) Assessment moderation: constructing the marks and constructing the students. *Assessment & Evaluation in Higher Education* 32 (6), 645-656.

- Orr, S. and Bloxham, S. (2013) Making judgements about students making work: Lecturers' assessment practices in art and design. *Arts and Humanities in Higher Education* 12 (2-3), 234-253.
- Orsmond, P., Merry, S. and Reiling, K. (2000) The Use of Student Derived Marking Criteria in Peer and Self-assessment. *Assessment & Evaluation in Higher Education* 25 (1), 23-38.
- Paley, J. (1997) Husserl, phenomenology and nursing. *Journal of Advanced Nursing* 26 (1), 187-193.
- Paley, V. G. (1990) *The boy who would be a helicopter: The use of storytelling in the classroom*. Cambridge M.A.: Harvard University Press.
- Parahoo, K. (2006) *Nursing Research Principles, Process and Issues*. Second edition. Hampshire: Palgrave MacMillan.
- Pejtersen, J. H., Feveile, H., Christensen, K. B. and Burr, H. (2011) Sickness absence associated with shared and open-plan offices — a national cross sectional questionnaire survey. *Scandinavian Journal of Work, Environment & Health* 37 (5), 376-382.
- Peter, I. *Today if you are not confused*. Abundant Empowerment. [http://www.abundantempowerment.com/index.php?option=com\\_content&view=article&id=1592](http://www.abundantempowerment.com/index.php?option=com_content&view=article&id=1592) Accessed 14 August.
- Pinot de Moira, A., Massey, C., Baird, J.-A. and Morrissy, M. (2002) Marking consistency over time. *Research in Education* 67 (1), 79.
- Polanyi, M. (1966) The logic of tacit inference. *Philosophy* 41 (155), 1-18.
- Polanyi, M. (2009) *The Tacit Dimension*. 02/01 edition. Vol. 41. The University of Chicago Press, Ltd, London: The Doubleday Broadway Publishing Group.
- Popham, W. J. (1997) What's Wrong--and What's Right--with Rubrics. *Educational leadership* 55 (2), 72-75.
- Price, M. (2005) Assessment standards: the role of communities of practice and the scholarship of assessment. *Assessment & Evaluation in Higher Education* 30 (3), 215-230.
- Price, M., Handley, K., Millar, J. and O'Donovan, B. (2010) Feedback: all that effort, but what is the effect? *Assessment & Evaluation in Higher Education* 35 (3), 277-289.
- Price, M. and Rust, C. (1999) The experience of introducing a common criteria assessment grid across an academic department. *Quality in Higher Education* 5 (2), 133-144.

- QAA. (2012) *Understanding assessment: its role in safeguarding academic quality and standards in Higher Education. A guide for early career staff.* (978 1 84979 390 2). Quality Assurance Agency for Higher Education. Gloucester: Quality Assurance Agency for Higher Education.  
<http://www.qaa.ac.uk/Publications/InformationAndGuidance/Pages/understanding-assessment.aspx>
- QAA. (2013) *UK Quality Code for Higher Education Part B.* Quality Assurance Agency for Higher Education. Gloucester: QAA.  
<http://www.qaa.ac.uk/en/Publications/Documents/understanding-assessment.pdf> Accessed December 15.
- QAA. (2016) *Subject Benchmark Statements Health Studies.* Quality Assurance Agency for Higher Education. Gloucester: QAA.  
<http://www.qaa.ac.uk/en/Publications/Documents/SBS-Health-Studies-16.pdf>
- Read, B., Francis, B. and Robson, J. (2005) Gender, 'bias', assessment and feedback: analyzing the written assessment of undergraduate history essays. *Assessment & Evaluation in Higher Education* 30 (3), 241-260.
- Reddy, Y. M. and Andrade, H. (2010) A review of rubric use in higher education. *Assessment & Evaluation in Higher Education* 35 (4), 435-448.
- Rezaei, A. R. and Lovorn, M. (2010) Reliability and validity of rubrics for assessment through writing. *Assessing Writing* 15 (1), 18-39.
- Rowe, N. and Martin, R. (2014) Dancing onto the page: Crossing an academic borderland. *Waikato Journal of Education*, 19 (2), 12.
- Rust, C., Price, M. and O'Donovan, B. (2003) Improving Students' Learning by Developing their Understanding of Assessment Criteria and Processes. *Assessment & Evaluation in Higher Education* 28 (2), 147-164.
- Sadala, M. L. A. and Adorno, R. d. C. F. (2002) Phenomenology as a method to investigate the experience lived: a perspective from Husserl and Merleau Ponty's thought. *Journal of Advanced Nursing* 37 (3), 282-293.
- Sadler, D. R. (1987) Specifying and Promulgating Achievement Standards. *Oxford Review of Education* 13 (2), 191-209.
- Sadler, D. R. (1989) Formative assessment and the design of instructional systems. *Instructional science* 18 (2), 119-144.

- Sadler, D. R. (2005) Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education* 30 (2), 175-194.
- Sadler, D. R. (2007) Perils in the meticulous specification of goals and assessment criteria. *Assessment in Education* 14 (3), 387-392.
- Sadler, D. R. (2009a) Fidelity as a precondition for integrity in grading academic achievement. *Assessment & Evaluation in Higher Education* 35 (6), 727-743.
- Sadler, D. R. (2009b) Grade integrity and the representation of academic achievement. *Studies in Higher Education* 34 (7), 807-826.
- Sadler, D. R. (2009c) Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education* 34 (2), 159-179.
- Sadler, D. R. (2009d) Transforming holistic assessment and grading into a vehicle for complex learning. *Assessment, learning and judgement in higher education*. Springer. 1-19.
- Sadler, D. R. (2013) Assuring academic achievement standards: from moderation to calibration. *Assessment in Education: Principles, Policy & Practice* 20 (1), 5-19.
- Sadler, D. R. (2014) The futility of attempting to codify academic achievement standards. *Higher Education* 67 (3), 273-288.
- Sambell, K., McDowell, L. and Montgomery, C. (2012) Assessment for learning in higher education. Abingdon, Oxon: Routledge. 172.
- Saunders, M. N. K. and Davis, S. M. (1998) The use of assessment criteria to ensure consistency of marking: some implications for good practice. *Quality Assurance in Education* 6, 162-171.
- Saywitz, K. J., Geiselman, R. E. and Bornstein, G. K. (1992) Effects of cognitive interviewing and practice on children's recall performance. *Journal of Applied Psychology* 77 (5), 744-756.
- Schön, D. A. (1987) Educating the reflective practitioner: Toward a new design for teaching and learning in the professions. *San Francisco*.
- Schön, D. A. (1995) Knowing-in-action: The new scholarship requires a new epistemology. *Change: The Magazine of Higher Learning* 27 (6), 27-34.
- Sellbjer, S. (2015) Meaning in constant flow – university teachers' understanding of examination tasks. *Assessment & Evaluation in Higher Education*, 1-13.

- Servage, L. (2009) The Scholarship of Teaching and Learning and the Neo-Liberalization of Higher Education: Constructing the "Entrepreneurial Learner". *The Canadian Journal of Higher Education* 39 (2), 25.
- Silvestri, L. and Oescher, J. (2006) Using Rubrics to Increase the Reliability of Assessment in Health Classes. *International Electronic Journal of Health Education* 9, 25-30.
- Smith, C. and Boyd, P. (2012) Becoming an academic: the reconstruction of identity by recently appointed lecturers in nursing, midwifery and the allied health professions. *Innovations in Education and Teaching International* 49 (1), 63-72.
- Smith, C. D., Worsfold, K., Davies, L., Fisher, R. and McPhail, R. (2013) Assessment literacy and student learning: the case for explicitly developing students 'assessment literacy'. *Assessment & Evaluation in Higher Education* 38 (1), 44-60.
- Smith, D. W. (2011) *Phenomenology*. The Stanford Encyclopedia of Philosophy:  
<http://plato.stanford.edu/archives/fall2011/entries/phenomenology/>  
 Accessed 25 January.
- Smith, J. A., Flowers, P. and Larkin, M. (2009) *Interpretative phenomenological analysis: Theory, method and research*. SAGE Publications Limited.
- Soliman, I. and Soliman, H. (1997) Academic Workload and Quality. *Assessment & Evaluation in Higher Education* 22 (2), 135-157.
- Sonne, M., Villalta, D. L. and Andrews, D. M. (2012) Development and evaluation of an office ergonomic risk checklist: ROSA–Rapid office strain assessment. *Applied ergonomics* 43 (1), 98-108.
- Taras, M. (2002) Using Assessment for Learning and Learning from Assessment. *Assessment & Evaluation in Higher Education* 27 (6), 501-510.
- Taras, M. (2006) Do unto others or not: equity in feedback for undergraduates. *Assessment & Evaluation in Higher Education* 31 (3), 365 - 377.
- Thomas, S. P. (2005) Through the lens of Merleau-Ponty: advancing the phenomenological approach to nursing research. *Nursing Philosophy* 6 (1), 63-76.
- Thompson, G. (2001) Interaction in academic writing: Learning to argue with the reader. *Applied linguistics* 22 (1), 58-78.

- Thomson, P. and Gunter, H. (2011) Inside, outside, upside down: the fluidity of academic researcher 'identity' in working with/in school. *International Journal of Research & Method in Education* 34 (1), 17-30.
- Tomas, C. (2013) Marking and feedback provision on essay-based coursework: a process perspective. *Assessment & Evaluation in Higher Education* 39 (5), 611-624.
- Tomas, C. (2014) Marking and feedback provision on essay-based coursework: A process perspective. *Assessment & Evaluation in Higher Education* 39 (5), 611-624.
- Tourangeau, R. (1999) Remembering What Happened: Memory Errors and Survey Reporting. In Stone, A. A. B., Christine A. Jobe, Jared B. Kutzman, Howard S. Cain, Virginia S. (editor) *The Science of Self-report: Implications for Research and Practice*. Vol. 1. Psychology Press. 392.
- Tourangeau, R., Rips, L. J. and Rasinski, K. (2000) *The psychology of survey response*. Cambridge University Press.
- Trede, F. and Smith, M. (2014) Workplace educators' interpretations of their assessment practices: a view through a critical practice lens. *Assessment & Evaluation in Higher Education* 39 (2), 154-167.
- Trowler, P. and Knight, P. T. (2000) Coming to Know in Higher Education: Theorising faculty entry to new work contexts. *Higher Education Research & Development* 19 (1), 27-42.
- Tuck, J. (2015) 'That ain't going to get you a professorship': discourses of writing and the positioning of academics' work with student writers in UK higher education. *Studies in Higher Education*, 1-15.
- UoB (2015) *Guide to External Examining For Taught Programmes 2015-2016*. University of Bradford.
- Van Manen, M. (1990) *Researching Live Experience: Human science for an action sensitive pedagogy*. Ontario, Canada: The State University of New York Press.
- Walker, M. (2009) An investigation into written comments on assignments: do students find them usable? *Assessment & Evaluation in Higher Education* 34 (1), 67-78.
- Ward, D. J., Furber, C., Tierney, S. and Swallow, V. (2013) Using framework analysis in nursing research: a worked example. *Journal of advanced nursing* 69 (11), 2423-2431.



- Weaver, S. D. and Gahegan, M. (2007) Constructing, visualizing and analyzing a digital footprint. *Geographical Review* 97 (3), 324-350.
- Webb, C. (1992) The use of the first person in academic writing: objectivity, language and gatekeeping. *Journal of Advanced Nursing* 17 (6), 747-752.
- Wenger, E. (1998) Communities of practice: Learning as a social system. *Systems thinker* 9 (5), 2-3.
- Wenger, E. (2011) Communities of practice: A brief introduction. 7. <https://scholarsbank.uoregon.edu/xmlui/bitstream/handle/1794/11736/a%20brief%20introduction%20to%20cop.pdf?sequence=1> Accessed 24/05/2016
- Wenger, E. C. and Snyder, W. M. (2000) Communities of practice: The organizational frontier. *Harvard business review* 78 (1), 7.
- William, D. (2008) What should Educational Research Do, and How Should it do it? *Educational Researcher* 37 (7), 7.
- Wimshurst, K. and Manning, M. (2012) Feed-forward assessment, exemplars and peer marking: evidence of efficacy. *Assessment & Evaluation in Higher Education*, 1-15.
- Wojnar, D. M. and Swanson, K. M. (2007) Phenomenology: An Exploration. *Journal of Holistic Nursing* 25 (3), 172-180.
- Wyatt-Smith, C., Klenowski, V. and Gunn, S. (2010) The centrality of teachers' judgement practice in assessment: a study of standards in moderation. *Assessment in Education: Principles, Policy & Practice* 17 (1), 59-75.
- Yorke, M. (2010) How Finely Grained Does Summative Assessment Need to Be? *Studies in Higher Education* 35 (6), 677-689.
- Yorke, M. (2011) Summative Assessment: Dealing with the "Measurement Fallacy". *Studies in Higher Education* 36 (3), 251-273.
- Yorke, M., Barnett, G., Bridges, P., Evanson, P., Haines, C., Jenkins, D., Knight, P., Scurry, D., Stowell, M. and Woolf, H. (2002) Does grading method influence honours degree classification? *Assessment & Evaluation in Higher Education* 27 (3), 269-279.
- Yorke, M., Bridges, P. and Woolf, H. (2000) Mark distributions and marking practices in uk higher education some challenging issues. *Active learning in higher education* 1 (1), 7-27.
- Yosso, T. J. (2005) Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethnicity and Education* 8 (1), 69-91.

# APPENDICES

## 1. Ethics Approval



**UNIVERSITY OF  
BRADFORD**  
MAKING KNOWLEDGE WORK™

**Research and Knowledge Transfer Support**  
Richmond Road | Bradford | West Yorkshire | BD7 1DP | UK  
Tel: +44 (0)1274 236000  
Fax: +44 (0)1274 236090  
[www.bradford.ac.uk/research-and-business](http://www.bradford.ac.uk/research-and-business)

Fiona Meddings  
Lecturer  
Midwifery & Reproductive Health  
School of Health Studies

12<sup>th</sup> December 2013

**Ethics Application:** E323 Hughes/Meddings  
**Application Title:** Demystifying marking and assessment processes: unmasking  
assessment by revealing the concept of an 'expert' marker

Your Ethics Application has now been reviewed, with the Chair of the Research Ethics Panel completing a consideration of the review report.

I am pleased to inform you that the Chair of the Panel has confirmed approval of this study, with no further ethical scrutiny.

Please add a sentence onto any material you share with participants confirming that ethics approval has been granted to you by the Chair of the Humanities, Social and Health Sciences Research Ethics Panel at the University of Bradford on 12<sup>th</sup> December 2014.

Kind Regards



Clare Beckett  
Chair, Humanities, Social and Health Sciences Research Ethics Panel  
University of Bradford  
T: 01274 233112  
E: [p.f.ali@bradford.ac.uk](mailto:p.f.ali@bradford.ac.uk)



Date 25102014

Ref. 2014-5/HWB/HSC/EXTERNAL/2

Dear Fiona Meddings

This letter relates to your proposal: Demystifying marking and assessment processes: unmasking assessment by revealing the concept of an 'expert' marker

This proposal was submitted to me as Chair of the Faculty Research Ethics Committee for an opinion on its suitability to be undertaken in Sheffield Hallam University. I note you have Ethics approval from the University of Huddersfield. As such, I am happy to grant permission for the project to go ahead without approval from the FREC.

The documents I examined have been combined into one PDF named Meddings 2014.pdf I have saved and filed it electronically.

I have saved these as PDF documents and filed them electronically.

Good luck with your project.

Yours sincerely



Peter Allmark  
Chair Faculty Research Ethics Committee

Faculty of Health and Well-being  
Sheffield Hallam University  
32 Collegiate Crescent  
S10 2BP

0114 225 5727  
[p.allmark@shu.ac.uk](mailto:p.allmark@shu.ac.uk)



University Research Ethics Committee  
Research Business and Innovation  
Frenchay Campus  
Bristol BS16 1QY

Email [res.admin@uwe.ac.uk](mailto:res.admin@uwe.ac.uk)  
Telephone +44 (0)117 32 82872

UWE REC REF No: UREC14-15/02

30<sup>th</sup> October 2014

Fiona Meddings  
Researcher and PhD Student  
Lecturer, Division of Midwifery & Reproductive Health  
University of Bradford  
School of Health Studies  
Richmond Road  
Bradford  
BD50BB

Dear Fiona

**Application title: Demystifying marking and Assessment Processes: unmasking assessment by revealing the concept of an 'expert' marker**

Your ethics application was considered by the University Research Ethics Committee and, based on the information provided, has been given ethical approval to proceed.

The following standard conditions also apply to all research given ethical approval by a UWE Research Ethics Committee:

1. You must notify the relevant UWE Research Ethics Committee in advance if you wish to make significant amendments to the original application: these include any changes to the study protocol which have an ethical dimension. Please note that any changes approved by an external research ethics committee must also be communicated to the relevant UWE committee.
2. You must notify the University Research Ethics Committee if you terminate your research before completion;
3. You must notify the University Research Ethics Committee if there are any serious events or developments in the research that have an ethical dimension;
4. Any changes to the study protocol, which have an ethical dimension, will need to be approved by the research ethics committee. You should send details of any such amendments to the relevant committee with an explanation of the reason for the proposed changes. Any changes approved

UREC/FREC Standard Approval Letter

Version 1 1/8/2013



Ms Fiona Meddings  
Lecturer  
Midwifery & Reproductive Health  
School of Health Studies  
University of Bradford  
Bradford  
BD7 1DP

**FACULTY OF HEALTH  
AND SOCIAL CARE**  
T: 01482 464680  
E: [j.dyson@hull.ac.uk](mailto:j.dyson@hull.ac.uk)

**OUR REF: 165**  
03 November 2014

Dear Fiona

**Re: Demystifying marking and assessment processes: unmasking assessment by revealing the concept of an 'expert' marker**

Thank you for your submission to the Faculty of Health & Social Care Research Ethics Committee, including ethics approval from the University of Bradford Humanities, Social and Health Sciences Research Ethics Panel.

Given the information you have provided, I am able to grant Chair's approval as per this Committee's Terms of Reference.

In order to request access to potential participants please contact the Heads of Department as follows:

Professor Julie Jomeen, Dept. Midwifery & Child Health and Psychological Health & Wellbeing ([J.Jomeen@hull.ac.uk](mailto:J.Jomeen@hull.ac.uk))

Professor Mark Hayter, Dept Nursing ([M.Hayter@hull.ac.uk](mailto:M.Hayter@hull.ac.uk))

Dr Mary Laurenson, Acting HOD Health Technology & Peri-operative Practice ([M.C.Laurenson@hull.ac.uk](mailto:M.C.Laurenson@hull.ac.uk))

I wish you every success with your research.

Yours sincerely



Dr Judith Dyson  
Chair, Research Ethics Committee

cc: file

## 2. Interviewee Information Sheet



### PARTICIPANT INFORMATION LEAFLET

#### **DEMYSTIFYING MARKING AND ASSESSMENT PROCESSES. A PHENOMENOLOGICAL ANALYSIS OF MARKING AND ASSESSMENT PRACTICES OF NOVICE AND EXPERIENCED MARKERS**

Researcher: Fiona Meddings, PhD student

Principle Supervisor: Peter Hughes, Centre for Educational Development, University of Bradford.

Thank you for considering taking part in this research study. This leaflet explains more about the research.

Before you decide I would like you to understand why the research is being done and what it would involve for you. I will go through the information sheet with you and answer any questions you have. I suggest this should take about 10 -15 minutes. Talk to others about this study if you wish. Whether or not you participate in this study, it will have no effect on your job, role requirements or status. Please ask if there is anything that is not clear.

As a PhD student my research is an exploration of the process undertaken by lecturers when marking and assessing. I want to examine the process of marking from the perspectives of lecturers to identify ways in which support could be made available for novice markers and those new to higher education.

You are invited to take part in this research study. Before you decide, it is important for you to understand why the research is being carried out. Please take time to read the following information carefully and decide if you wish to participate in this study.

#### **What is the purpose of this study?**

The purpose of this study is to explore marking and assessment practices undertaken by lecturers in higher education. To do this I intend to interview in one of two ways either whilst you are actually completing a marking task or a post marking reflection on the process. Interviews will focus on your experience of marking and assessing written student work. Findings from this study will inform understanding about how lecturers mark and assess student work, providing in-depth reflection on the actual act of marking. This study will develop conceptual understandings about the process undertaken by lecturers to provide a mark or grade for a student

submission for assessment. This has the potential to provide information to influence the development of novice lecturers, by incorporating findings in higher education development programmes.

### **Why have I been chosen?**

You have been chosen because you are a lecturer in a health studies subject supporting and marking or grading student work at undergraduate or postgraduate level.

### **Do I have to take part?**

No it is entirely up to you to decide whether or not to take part. There will be no adverse consequences for you if you decide not to take part. I will describe the study and go through the information sheet. I will ask for your formal, consent prior to your participation in the study. This will initially be collected via email, however at the time of us meeting for the face to face interview I will check with you that you are still happy to go ahead. You can withdraw from the study at any time without giving a reason. Whether or not you participate in this study, it will have no effect on your job, role requirements or status.

### **What do I have to do?**

You will be asked to agree to one interview which will either be during or after an authentic marking task. The interviews will last approximately 90 minutes each. However they may be shorter or indeed a little longer depending upon which interview technique is being used. You will not need to allocate any more than 120 minutes to this activity. It is envisaged that the interviews will take place following submission of student written essay work, for which you have responsibility to mark or grade.

### **What are the possible risks of the research?**

I do not anticipate that taking part in this research will incur any risks.

### **What are the possible benefits of the research?**

Whilst the information may not be of direct benefit to you personally, the information I gather will help to identify information that could successfully improve marking and assessing amongst novice or inexperienced markers. The research is intended ultimately to be of benefit to lecturers working in higher education with responsibilities for marking and assessing student work. The outcomes of the research will potentially be of benefit to those developing training and education for higher education professionals.

### **Will my involvement be confidential?**

Yes. I will follow ethical and legal practices in respect of confidentiality. The interviews will be recorded and transcribed to assist accurate analysis



and will be stored securely with an anonymous code. Notes on the observations will be stored in the same way. The custodian of the data will be the Chief Investigator of this research project, the research supervisor. Typed transcripts and field notes from the study will be stored anonymously. All data will be handled, processed, stored and destroyed by the researcher in compliance with the Data Protection Act 1988. All data collected from you will be anonymised and given a unique identifier. Only the researcher will have access to the data, which will be kept safe, stored in a locked filing cabinet or on a password protected computer. All personal data (i.e. name, address, telephone number) will be destroyed after three years.

### **Who is organising and funding the research?**

The researcher is based at the University of Bradford, School of Health Studies and is accordingly supported and regulated by the University.

### **Who has reviewed the study?**

All research is reviewed by the University research ethics committee to ensure that your interests as a participant are protected. Ethics approval has been given for this study by the Humanities, Social Sciences and Health Studies Research Ethics Panel at the University of Bradford on 12th December 2013, E.323.

### **Further Information**

If you have any queries regarding this project please contact the researcher:

Fiona Meddings

Researcher and PhD Student  
Lecturer, Division of Midwifery & Reproductive Health  
University of Bradford  
School of Health Studies  
Richmond Road  
Bradford  
BD50BB  
Tel: 01274 236479  
Email: [F.S.Meddings@Bradford.ac.uk](mailto:F.S.Meddings@Bradford.ac.uk)



## 4. Recruitment Protocol



An invitation to take part in a research study: ¶

### “Demystifying Marking and Assessment Processes” ¶

¶

Researcher: ··Fiona Meddings, PhD student. ¶

Have you ever had time to stop and think about your marking, about how you produce the mark, grade and feedback you do? ··Do you think you would be able to explain the process you undertake to someone who has never marked before? ¶

My PhD research is examining the process of marking from the perspectives of lecturers to identify ways in which support could be made available for novice markers and those new to higher education. ¶

I am particularly interested in recruiting early career lecturers who are in their first academic post and have completed less than 2 years in academia. ¶

I need **YOU**, whether you are experienced or new to marking or grading student work. You are invited to take part in this research study. ··Ethics approval has been given for this study by the Humanities, Social Sciences and Health Studies Research Ethics Panel at the University of Bradford on 12th December 2013. ··It has also received ethical approval from the Chair of the Faculty of Health and

external institution. ¶

If you would like to lend your support, contact me by email so that you can find out more and for us to arrange a convenient time and location for an interview. ··All my data collection must be completed by 31 January 2015, so should fit within with any Semester 1 marking schedules. ¶

Email: ··[F.S.Meddings@Bradford.ac.uk](mailto:F.S.Meddings@Bradford.ac.uk) ¶

¶

## 5. Consent Forms



### COGNITIVE-INTERVIEW-CONSENT-FORM¶

#### **Study: Demystifying Marking and Assessment Processes. A phenomenological analysis of marking and assessment practices of novice and experienced markers¶**

Thank you for agreeing to take part in this study which is exploring marking and assessment practices undertaken by lecturers in higher education. This interview will require you to reflect on a marking event in the past and it will focus on your experience of marking and assessing student written work.¶

The interview should take no longer than 90 minutes though our discussion may conclude sooner. With your permission a digital device will be used to record your interview. I may also make notes during the course of the interview. All records will be stored securely and anonymity will be maintained. No personal data will be stored with interview transcripts or researcher notes.¶

I agree to participate in the interview and for anonymised data to be used for research and reporting purposes.¶

Print Name: \_\_\_\_\_¶

Signature: \_\_\_\_\_¶

Date of Interview: \_\_\_\_\_¶

Signature of interviewer: \_\_\_\_\_¶

Thank you for taking part in this research project.¶

Please complete and return this form by email to [F.S.Meddings@Bradford.ac.uk](mailto:F.S.Meddings@Bradford.ac.uk) prior to your agreed interview date.¶

PROTOCOL ANALYSIS INTERVIEW CONSENT FORM

**Study: Demystifying Marking and Assessment Processes – A  
phenomenological analysis of marking and assessment practices  
of novice and experienced markers**

Thank you for agreeing to take part in this study which is exploring marking and assessment practices undertaken by lecturers in higher education. This interview will require you to complete an authentic marking event whilst the researcher asks you to speak your thoughts out loud. This type of interview is called a 'Talk Aloud' interview and will enable the recording of the process of marking as it occurs. A short post-marking event interview will focus on your experience of marking and assessing student written work, as well as giving you opportunity to debrief after being recorded and observed undertaking marking or grading.

The interview should take no longer than 120 minutes, this will allow for you to spend as much time as you require for marking or grading the student essay. However the marking process and our discussion may conclude sooner. With your permission a digital device will be used to record your interview. I may also make notes during the course of the interview. All records will be stored securely and anonymity will be maintained. No personal data will be stored with interview transcripts or researcher notes.

I agree to participate in the interview and for anonymised data to be used for research and reporting purposes.

Print Name: \_\_\_\_\_

Signature: \_\_\_\_\_

Date of Interview: \_\_\_\_\_

Signature of interviewer: \_\_\_\_\_

Thank you for taking part in this research project.

Please complete and return this form by email to [F.S.Meddings@Bradford.ac.uk](mailto:F.S.Meddings@Bradford.ac.uk) prior to your agreed interview date.

## 6.CI Protocol

### COGNITIVE INTERVIEW PROTOCOL

General **introduction** to topic –

Demystifying marking and assessing (grading) practices. A

phenomenological analysis of marking and assessment practices of novice and experienced markers.

### **Introduction to research project –**

Aim:

To investigate the process of marking from the perspective of the marker, establishing individual marking practices in order to develop the proficiency of novice markers.

Objectives:

1. Examine assessing and marking practices – (what do markers assess/mark?, When, How, Why)
2. Identify what marking practices are undertaken, to develop a concept of 'expert' marking
3. Explore novice markers thoughts on marking/assessing.
4. Explore cognitive processes and extraneous influences on marking/assessment practice

Research Questions

- o How do Lecturers mark, how do they know how to mark, how do they develop their practice (or expertise)?
1. What thought or cognitive processes occur in lecturers during the act of marking?
  2. What do lecturers believe they do when they are undertaking marking?
  3. What exists to ensure equity or consistency in uni-professional or intra-professional (including inter-rater reliability) or inter-disciplinary marking?

- o *Recording of the interview*
- o Assure re confidentiality – but excerpts from answers may be used to elaborate or illuminate the discussion in my report.
- o Sign a consent form

## **COGNITIVE INTERVIEW PROTOCOL**

### *Questions Schedule and Probes*

#### **General Scene setting**

In the first instance the focus is on a single module, thinking about how you reached decisions in relation to providing the final mark or grade awarded to the student (prior to any moderation processes).

#### **General Probe**

- o How do you go about marking or grading a student paper?
- o Try to reflect on your most recent marking
- o How did you arrive at the final grade for the paper?
- o What went on in your mind when you were reading the paper?
- o Was it easy or difficult to arrive at a final grade? Why?

#### **Comprehension Probes**

- o Tell me in your own words what marking, assessment and grading mean to you
- o What do you understand by marking, assessing or grading?
- o Are these one and the same or different?
- o Can you tell me in your own words what goes through your mind when you are marking and assessing a student paper

#### **Retrieval Probes**

- o What time period were you thinking about when you did the marking?
- o When was this and what module was it?
- o Were you marking inter professional or uni-professional?
- o How did you remember what you were thinking?
- o Was there anything specific about what you were doing at the time?
- o How did you go about calculating the final mark and grade awarded?
- o Did you manipulate it in any way?
- o If so why?

### **Judgement Probes**

- o How sure are you of the final mark and grade you awarded?
- o What did you think about when deciding how to award the final mark or grade
- o How accurate would you say the final mark or grade you awarded was?

### **Response Probes**

- o How easy or difficult was it to select a category from the options provided on the mark sheet which related to the work you were marking/assessing/grading?
- o Are there any specific categories missing from the options provided that might have helped you?
- o What is it that is missing?
- o Why did you select

### **Sensitivity probes**

- o How do you feel about discussing your approach to marking and assessing?
- o Did you find these questions difficult to answer? If so why do you think that is?
- o Do you find these questions to be intrusive? If so why?

Rationale: "building an authentic story of assessment".

### **Expansive Probing**

- o Reorienting probes – directs verbose response back to the purpose of marking
- o Can you give me a clear example of X?
- o Elaborative probes – takes a side track
- o Tell me more about why X?



# 7. PA Interview Schedule

## PROTOCOL ANALYSIS INTERVIEW GUIDELINES

### General introduction to topic –

*Demystifying marking and assessing (grading) practices. A phenomenological analysis of marking and assessment practices of novice and experienced markers.*

### Introduction to research project –

*Aim:*

To investigate the process of marking from the perspective of the marker, establishing individual marking practices in order to develop the proficiency of novice markers.

*Objectives:*

1. Examine assessing and marking practices – (what do markers or grader do? When, How, Why)
2. Identify what marking practices are undertaken, to develop a concept of 'expert' marking
3. Explore novice markers thoughts on marking/assessing.
4. Explore cognitive processes and extraneous influences on marking/assessment practice

### Research Questions

- How do Lecturers mark, how do they know how to mark, how do they develop their practice (or expertise)?
  1. What thought or cognitive processes occur in lecturers during the act of marking? Protocol Analysis
  2. What do lecturers believe they do when they are undertaking marking?
  3. What exists to ensure equity or consistency in uni-professional or intra-professional (including inter-rater reliability) or inter-disciplinary marking?
- Recording of the interview
- Assure re confidentiality – but excerpts from answers may be used to elaborate or illuminate the discussion in my report.
- Sign a consent form

## **Introduction to Protocol Analysis**

Protocol analysis (PA) conceived by Ericsson and Simon (1993) to gather information about the thought processes occurring during the performance of a behaviour or task. During a PA interview the lecturer will perform an authentic marking or grading task, entailing making an evaluation of a student essay for assessment. Capturing marking or grading in action is an appropriate way to highlight experiences. Lecturers will be able to verbalise what they are going, at the point they are carrying out the activity.

All utterances from the lecturer will be recorded. They are implored to 'Talk Aloud' in essence to speak all the thoughts that are occurring whilst undertaking the task.

## **Completing the interview**

A short debrief covering the following aspects:

- **How lecturer felt?**
- **Any difficulties?**
- **Anything that surprised them?**
- **What else could be done to help them verbalise?**

## 8. PA Training

### *Training on pa (script)*

The warm up sessions are introduced to assist the lecturer in becoming familiar with the practice of talking or verbalising their thoughts. This practice is required as it is not a normal everyday process to have to utter each thought that is going on inside our minds.

“I’m interested in what you think about when you are marking or grading the student essay. In order for me to do this I am going to ask you to THINK ALOUD as you work with the student paper. What I mean is that I want you to tell me EVERYTHING you are thinking from the time you first see the essay until you have completed your marking or grading including writing student feedback. I would like you to talk to aloud CONSTANTLY. I don’t want you to explain to me what you are doing. I want you to act as if you were in the room alone and speaking to yourself about the student paper in front of you. It is very important for you to keep talking; so if you are quiet for a long period of time I will ask you to talk. Do you understand what I want you to do? Great now we will begin with some practice problems. *First I want you to multiply these two numbers in your head and tell me what you are thinking as you get an answer (not whether the answer is right or wrong)”*

***Give a card with this question:***

***“What is the result of multiplying 24 x 36”***

“Good now I want to see how much you can remember about what you were thinking from the time you read the question until you gave the answer. I am interested in what you actually can REMEMBER rather than what you think you must have thought. If possible tell me about your memories in the sequence they occurred when working on the question. Please tell me if you are uncertain about any of your memories. I don’t want you to work on solving the problem again, just report all that you can remember thinking about when answering the question. Now tell me what you remember. Great. Now I will give you two more practice problems before we continue with you commencing marking or grading the student essay. I want you to do the same thing for each of these problems. I want you to think aloud as before as you think about the questions and after you have answered it I will ask you to report on all that you can remember about your thinking. Any questions? Here’s your next problem.

***Give a card with this question:***

***“How many windows are there in your parent’s/in-laws/best friend’s house?”***

Now tell me all that you can remember about your thinking.  
Good now here's another practice problem. Please think aloud as you try to answer it. There is no need to keep count, I will keep track for you.

***Give a card with this question:***

***"Name 20 animals"***

***Now tell me all that you can remember about your thinking."***

***(Ericsson and Simon 1993: :378)***