



A study of the PhD examination: process, attributes and outcomes

Gillian Houston
Green Templeton College
Department of Education

Thesis submitted to the University of Oxford for the degree of DPhil
1st October 2018

Contents

	Page
Abstract	5
Acknowledgements	6
Chapter 1: The evolved PhD and the context for the study	
1 The evolved PhD	7
2 Context for the study	8
3 Rationale	15
4 Research questions and study boundaries	17
5 Thesis content and structure: summary	18
Chapter 2: The impact of policy changes on assessment in higher education	
1 Introduction	20
2 Policy changes designed to influence academic practice	21
3 New ways of thinking about teaching, learning and assessment	22
4 The impact of the doctoral qualification descriptor	24
5 Peer review of doctoral assessment	26
6 The student voice	28
7 Intended student learning outcomes (ILOs) and their feasibility in the doctorate	30
8 The challenge of defining and assessing 'transferable' skills	32
9 PhD standards and subject communities	34
10 Conclusion	37
Chapter 3: The examination and its relationship to the modern PhD	
1 Introduction	39
2 Empirical studies	40
3 The examination's fitness for purpose in the modern context	61
4 Conclusion	74
Chapter 4: Theoretical approach, research design and methods	
1 Introduction	76
2 Theoretical perspective	76
3 Qualitative data generation	78
4 Case study methodology	79
5 Dimensions of the study	83
6 Ethical considerations	84
7 Access challenges	85
8 Sampling	89
9 Research design	91
10 Research methods	93
11 The cases: sequence and procedures	95
12 Viva observations	97
13 Interviews	99
14 Data analysis	101
15 Conclusion	102

Chapter 5: The PhD examination process	
1 Introduction	104
2 The examination process: a continuum of judgement	105
3 The role of the thesis	108
4 The purposes of the viva	118
5 The relative importance of the thesis and the viva and the relationship between them	134
6 Conclusion	145
Chapter 6: Attributes sought by examiners in the thesis and the candidate	
1 Introduction	147
2 Originality (or, contribution to knowledge)	149
3 Publishability	158
4 Research competence	166
5 Intellectual rigour	180
6 Conclusion	185
Chapter 7: Examination outcomes: diversity versus comparability, and grading	
1 Introduction	190
2 The range of attainment	193
3 Examination outcomes	200
4 Grading and distinctions	211
5 Conclusion	223
Chapter 8: Discussion of findings and critical reflection	225
The PhD examination as a holistic process	226
Attributes sought by examiners	230
Examination outcomes: how do they reflect candidate achievement?	232
Challenges and limitations	234
Conclusion	236
Chapter 9: Conclusions	
The PhD examination process	237
Attributes contributing to 'doctorateness'	238
Examiners' judgements of PhD outcomes	239
The potential impact of the study on the PhD examination	240
Future research	241
Conclusion	242
References	244

Tables	
1.1: Chronology of major external interventions in UK HE: 1986-2014	8
2.1: Doctoral qualification descriptor (QAA, 2008)	25
2.2: 2009 PRES report (HEA, 2009; Table 11: Thesis Examination 2009)	29
4.1: Age and subject profiles of PhD candidates: comparison with UK figures	90
4.2: Research methods and their purposes	94
4.3: Identification of roles	96
4.4: Summary of interviews	100
5.1: Approximate percentage weightings showing relative importance of the thesis and viva in the overall judgement	139
6.1: Comparison of attributes identified in this study with Nyquist's 'core competencies' (2002)	186
6.2: Attributes sought and where they are identified	188
7.1: Summary of common PhD examination outcomes and time periods within which corrections are completed	190
7.2: Examiners' and supervisors' experiences of candidates being awarded a lower degree or required to undertake major corrections	204
7.3: Examiners' and supervisors' experiences of candidates failing the examination	209
Figures	
4.1: The 'heart', components and boundaries of each case	82
4.2: Ten bounded cases: Observations and semi-structured interviews, November 2011 to May 2014	95
4.3: Positioning of individuals in each viva observation	98
5.1: Continuum of judgement	106
8.1: The thesis-viva nexus	227
8.2: Integrated judgement	231
Appendices:	
1: Case summaries A to J (including analytical framework)	255
2: Information sheet and consent form for participants	278
3: Exploratory letter to institutions	283
4: Interview schedules: supervisors and examiners; candidates; and independent chairs/convenors	286
5: Coding framework	293
6: Extracts from university research degree regulations	299

Abstract

The idiographic nature of the PhD examination raises challenges for assessment in higher education. While the examination follows different processes internationally, the submission and evaluation of a thesis or dissertation that demonstrates originality is common to all. In many countries, this is complemented by a viva (oral examination), which in the UK is held in private. The confidentiality of the viva contributes to examiners' judgements being considered highly individual, especially when compared with the cohort-based examination processes that prevail in the assessment of students on taught programmes.

The study aimed to explore UK PhD examiners' judgements through observing 10 vivas and conducting interviews with 43 different actors in the process. The theoretical framework combined a realist perspective with case study methodology and involved ten cases in different subjects at six institutions. The qualitative data generated shed light on: the examination process; the role of the thesis, the purposes of the viva and the relationship between them; the attributes sought by examiners; and how different examination outcomes reflect candidates' achievements.

Data support the idea of a 'continuum of judgement' that begins with initial evaluation of the thesis and ends after the viva. During the assessment period, examiners' individual and collective judgements develop until the jointly agreed outcome of the viva is reached. The study confirms the centrality of the thesis in the assessment, while demonstrating the value of the viva for examiners and candidates alike. The relationship between the thesis and viva is shown to be interdependent, the viva assuming greater importance if the thesis is borderline. Candidates particularly appreciated the 'rite of passage' signified by the viva.

Criteria employed by examiners to inform their judgements are multifaceted. Data imply that examiners are seeking professional and personal, as well as research attributes. Initial evaluation of the thesis determines if the candidate has 'done enough' for the award of PhD but in their detailed scrutiny examiners assess the candidate's deeper knowledge and understanding of their research and its relevance in the field, seeking originality or a contribution to knowledge. While examiners' principal focus is on the candidate's research achievements, the importance of qualities such as intellectual rigour, leadership and integrity, also emerge from the study. The viva allows examiners to probe these personal qualities in more depth than is possible by reading the thesis.

The wide range of achievement among PhD candidates above the 'threshold' judgement is acknowledged, although many believe a pass/fail outcome remains appropriate. Reasons for rejecting a grading scheme include the challenge of developing universally acceptable criteria, the potential to increase subjectivity and the wish to avoid grade inflation. Examination outcomes do not currently reflect the significance of professional and personal attributes in the judgement, and the variability in institutional regulations regarding recommendations open to examiners remains a concern, especially the flexibility inherent in the 'minor corrections' category, which nevertheless offers flexibility to accommodate candidates' individual circumstances.

The study concludes that the thesis remains central to examiners' judgements, that the viva and thesis are interdependent and the viva fulfils important purposes. The research has served to emphasise the complex nature of the PhD examination process and suggest that more explicit reference in the examination to the candidate's professional and personal attributes would better reflect modern expectations of the PhD.

Acknowledgements

My sincere thanks go to all who contributed to this study, especially the PhD candidates who generously allowed me to observe their vivas and everyone who gave up their time to be interviewed. I hope to have done justice to their wise words. All the candidates were hugely impressive and so obviously deserved their PhDs.

Particular thanks go to my supervisors, Ingrid Lunt and latterly, Velda Elliott. They faced challenges in helping me to develop my writing to fit the academic genre, but provided much valuable advice and supported me through difficult times. I much appreciate their kindness, hard work and intellectual rigour, which improved my work immeasurably.

My wonderful family and friends (in the UK and Canada) have also played a huge part in encouraging me to finish the project when I felt like giving up. I can't name everyone but they know who they are. My heartfelt thanks go to them for always being on my side and believing I could do it.

The thesis is dedicated to four exceptional people who I am so fortunate to have had in my life and whose unconditional love and support has been transformational: my parents, Joan and Alex Houston, David Buckingham and my aunt, Margaret Alicia Fenwick, who so generously contributed to my fees.

Chapter 1: The evolved PhD and the context for the study

The PhD is the monarch of the academic community. It is the very highest accomplishment that can be sought by students. It signals that its recipient is now ready, eligible, indeed obligated, to make the most dramatic shift in roles: from student to teacher, from apprentice to master, from novice or intern to independent scholar and leader. The PhD marks its holder as one charged to serve as a steward of the discipline and profession. (Walker, Golde, Jones, Conklin Bueschel and Hutchings, 2008:x)

1 The evolved PhD

Introduced in the UK by Oxford University in 1917, the PhD's transformative role and its standing as an academic qualification are undisputed. Yet over the past 20–30 years, it has undergone profound changes. Candidate numbers have increased significantly and the PhD¹ has been transformed from a degree primarily undertaken by those aiming to enter the 'academy' to a qualification with broader focus, taken by graduates envisaging a range of professional roles. The changes are primarily a result of initiatives intended to 'modernise' the PhD. They illustrate its evolving nature and purposes, and include: the introduction of structured doctoral training; team-based, interdisciplinary approaches to research; and an ever-increasing emphasis on 'employability'. Major changes in assessment practice in universities have brought about greater transparency and arguably increased objectivity. The unique format of the final examination, however, influenced by the PhD's individuality and its roots in research, remains largely unaffected and differs markedly from other forms of assessment. Given the evolution of the degree since it was introduced in the UK, it is timely to question whether the examination meets current expectations of fairness and validity and evaluates the full range of attributes now expected of PhD candidates.

The PhD's original purpose, which remained unchanged for many years, was to provide a qualification for scholars pursuing a form of research apprenticeship in middle age under the supervision of an experienced academic in the field, in the expectation they would become the next generation of '*stewards of the discipline*' (Noble, 1994; Walker et al., 2008:12–13).

Walker et al. describe the PhD's elite academic status and its role in the transition from apprentice to independent researcher, a role that remains relevant to PhD candidates, in an academic context and more broadly. They recognised that the apprenticeship model needed to be '*reclaimed*' (p.91), as the concept had largely been superseded by multiple sources of advice for PhD candidates, who now typically work alongside their peers, supervisors and

¹ The shortened form 'PhD' rather than the less common 'DPhil' is used throughout to represent the

others in an academic environment offering a broad-based network of support and training where they are encouraged to develop professional as well as academic skills. Research and scholarship, however, remain the core purpose of the PhD, with the concept of originality or a contribution to knowledge a key criterion in the examination.

2 Context for the study

After a long period of stability, the unprecedented external interventions that have characterised UK higher education since the 1980s have had a major impact on many aspects of doctoral education. Ensuing changes to the PhD have therefore been driven by external imperatives rather than organic development. Key milestones are summarised in Table 1.

External interventions affecting doctoral education	
1986	Research Assessment Exercise introduced, replaced by the Research Excellence Framework in 2014
1988	Committee of Vice-Chancellors and Principals proposed structured training and grading for doctorates
1992	Further and Higher Education Act passed, incorporating abolition of Council for National Academic Awards
1996	Publication of Review of Postgraduate Education, commissioned by the Higher Education Funding Council for England (HEFCE)
1997	Establishment of the Quality Assurance Agency for Higher Education (QAA), successor to the Academic Audit Unit and Higher Education Quality Council
1999	QAA published first edition of UK Code of Practice, section 1: Research degrees
2001	QAA published first edition of UK Framework for Higher Education Qualifications, including doctoral qualification descriptor
2002	Publication of Roberts report <i>Set for Success</i> and funding to support transferable skills development for doctoral candidates
2003	HEFCE issued consultation paper on improving threshold standards in research degrees
2003	HEFCE published threshold standards for research degrees following consultation
2004	QAA published revised UK Code of Practice, section 1: Research degrees
2005	HEFCE published first doctoral completion rates results, with link to quality of supervision
2007	HEFCE published second issues paper on completion rates, followed by findings on individual institutional completions
2008	QAA revised UK Framework for Higher Education Qualifications, with additional text supporting doctoral qualification descriptor
2012	Higher Education Commission report on postgraduate education
2014	QAA revised and re-named UK Code of Practice, section 1: Research degrees, to: UK Quality Code, section B11: Research degrees

Table 1.1: Chronology of major external interventions in UK HE: 1986–2014

These developments include some ground-breaking initiatives. The first was a radical proposal by the Committee of Vice-Chancellors and Principals (CVCP) in 1988 for 'taught courses' in doctorates, a precursor to structured training. The Further and Higher Education Act of 1992 removed the distinction between universities and polytechnics, opening up competition and increasing diversity among doctoral students. In 1997 the UK funding bodies established a Quality Assurance Agency for Higher Education (QAA), to manage a system of peer review in universities. QAA initiatives influencing the delivery and structure of PhDs include a Code of Practice on research degrees (QAA, 2011) and a Framework for Higher Education Qualifications (QAA, 2008) that included a doctoral qualification descriptor containing learning outcomes. Implementation of the Roberts recommendations (Roberts, 2002) transformed professional skills training for doctoral candidates, while in England, the 2003 'Threshold Standards' widely consulted on by HEFCE (2003) were integrated with the second research degrees Code of practice (QAA, 2004). In 2005 and for several years, with efficiency and value for money in mind, HEFCE monitored doctoral completion rates, (2005, 2007a, 2007b).

2.1 The global PhD and the knowledge economy

The PhD has common currency worldwide as a prestigious academic qualification. It carries with it expectations about what a 'doctor' can do and the qualities they possess (Nyquist, 2002). The international status and value of the PhD was identified some time before the 21st century 'globalisation' of economies and of higher education itself (Noble, 1994).

Subsequently it was inevitable that the world-wide expansion of student numbers, the internationalisation of higher education and recognition of the wider value of universities in the economy (Boud and Lee, 2009; Oxford Economics, 2017) would force governments to analyse the contribution of doctoral researchers to the 'knowledge economy' (OECD, 1996; HEFCE, 2003b) and, arguably, lead to the 'commodification' of graduates and a focus on their economic contribution (Lawrence and Sharma, 2002; Radder, 2010).

Economic re-evaluation of the doctorate made a practical reality of the statement that '*PhDs are as much in an international market as currency*' (Russell, 1988:12). Unsurprisingly, during the early years of the 21st century, the UK doctorate came under increased scrutiny by governments, funding bodies and other stakeholders who were mindful that the UK needed to remain internationally competitive in doctoral recruitment (Kemp, Archer, Gilligan and Humfrey, 2008; Schwab, 2014). This imperative was exemplified in a report by the Higher Education Commission (2012), which foregrounded the role of postgraduate education and

particularly research degrees in the UK's future sustainability and success. The creativity and new knowledge provided by doctoral graduates in the workplace is unchallenged, while the nature and purpose of the doctorate continues to be debated (Park, 2005; Denicolo and Park, 2010).

The focus on international competitiveness and the knowledge economy led to questions concerning the nature and purposes of the PhD. Was it still fit for purpose? Were all PhD graduates equipped with the knowledge and skills necessary for employment? Was it possible to identify a 'threshold standard' that would guarantee all doctoral graduates could contribute to the knowledge economy?

2.2 Growth, employability and completion rates

Successive government policies to democratise higher education and set ambitious targets for the proportions of young people studying at university have led to significantly increased numbers of PhD graduates. Earlier periods of modest growth were succeeded by more significant expansion, many embarking on doctorates seeking career enhancement (HEFCE, 2015), emphasising the importance of timely completion and employability. Between 2003-4 and 2011-12, postgraduate research student numbers in UK universities increased by 25%, from 87,415 to 109,065, compared with an 18% increase in postgraduate taught student numbers over the same period (UUK, 2013). The most recent figures available for the UK as a whole (UUK, 2017) suggest that of a total population of 113,175 research degree students (mainly doctoral), around 75% are studying full-time and 25% part-time.

This growth has brought about greater diversity. Candidates with different academic and personal backgrounds pursue the PhD for varied reasons and aspire to diverse careers, including higher education, where academic jobs are scarce. A recent UKRI initiative to introduce study loans for doctoral candidates recognises the critical importance of financial support for those embarking on doctoral study, especially self-funding students, and affirms the perception of the doctoral degree as a career- and life-enhancing factor.

The focus on 'employability' has increased universities' commitment to incorporating so-called 'transferable' skills, considered valuable by employers, at all degree levels. However, research degree graduates continue to secure high rates of employment. According to the Higher Education Statistics Agency, in 2014/15, 74% of full-time doctorate leavers were in

employment compared to 67% for first degree leavers (HESA, 2016). Even if some forms of employment at this early stage in their careers are not directly related to their discipline or qualification status, the high level of employability of doctoral graduates belies some of the most extreme employer criticisms. The long-standing success of doctoral graduates in the jobs market has not always been complemented by high completion rates in doctoral programmes, however.

In England, completion rates were the focus of a HEFCE (2003) initiative to improve threshold standards in research degrees, their rationale being clearly linked to the employability agenda. HEFCE suggested that universities were fulfilling the first purpose of research degrees, i.e. to produce scholarly work that contributed to the knowledge economy, but that they were underperforming in the development of *'research and other skills that will in many cases go far wider than the original research'* (HEFCE, 2003:3). A key strand of the 'Improving standards' project was to improve completion and HEFCE (2005) published an Issues paper clearly linking quality of supervision to research degree completion rates. Based on a 7-year completion period, HEFCE established a 'best estimate' of 'eventual' completion rates to be between 71% and 82% for full-time research degrees and between 34% and 62% for part-time degrees. Universities had the opportunity to adjust their completion rates (HEFCE, 2007a) before individual institutions' completions were re-published (HEFCE, 2007b).

The growing intervention by external agencies, including the focus on employability and completion rates, was concurrent with the increased structure in doctoral education promoted to a large degree by the UK Research Councils and also encouraged by other developments, especially the 'cohort' model. The so-called 'structured PhD' supported a multi-purpose doctorate, intended to equip candidates with a range of academic, professional and personal skills that simultaneously enhanced employability and encouraged timely completion.

2.3 The structured PhD

The government's commitment to graduate employability and the acquisition of professional skills has had a major influence on the PhD's structure. In the past 20 or so years, therefore, what is now known as 'the structured PhD' has become the prevalent model for UK doctoral programmes. This remodelling of the traditionally unstructured and highly individual degree emerged from three developments: the commitment of the UK Research Councils to 'transferable' skills; the 'Roberts' recommendations (Roberts, 2002) and associated funding;

and the increasing influence of 'professional' doctorates designed for candidates to undertake original research to advance professional practice.

The introduction of formal structure in the PhD was a major development. It provided the first opportunity to create student cohorts in the early years, a significant change given that previously the focus of doctoral study was on individual progress in research. Adoption of the cohort model encouraged PhD entrants to engage with skills development opportunities designed to increase employability. The introduction of doctoral training centres (CDTs or DTCs) by the UK Research Councils, beginning with the EPSRC in 2003 as part of a wider initiative to promote interdisciplinary research, provided a further opportunity for skills development (Lunt, McAlpine and Mills, 2014). Cohort-driven courses were introduced for all UK Research Council funded students. If funding allowed, other doctoral candidates could participate.

Initiatives to introduce greater consistency in doctoral training embraced discipline-specific models. Whether based in a laboratory, technical, clinical, text, artefact or performing arts context, PhD candidates can now expect to experience various forms of structured training, from research methods to professional skills, often with others in their cohort, within a graduate school or doctoral college (UKCGE, 2015), and occasionally externally to the university. Attendance may be negotiated in response to individual needs, taking account of the candidate's mode of study, background and career intentions.

The CVCP (1988:2) first proposed the introduction of '*taught course elements*' in PhD programmes more than a decade earlier than Roberts (2002), who identified the need for transferable skills development, one element of a wide-ranging project concerning the supply of STEM graduates in the workplace. At a time when the doctorate was being criticised for its narrowness and limited relevance to sectors outside higher education, and the quality of PhD entrants and graduates was a concern, Roberts argued that there were not enough opportunities for candidates to acquire skills valued by employers: '*In particular, there is insufficient access to training in interpersonal and communication skills, management and commercial awareness*' (p.111). He further suggested that giving candidates more control over opportunities for development could be one of several factors that would increase the attractiveness of PhD study.

Through the research councils, government allocated significant funding to implement the Roberts recommendations. The more structured PhD model that emerged from introducing skills training, originally aimed only at STEM subjects, is now experienced by the majority of UK candidates, who have opportunities to gain a range of transferable skills while learning how to conduct research independently.

Concurrently with implementation of the Roberts recommendations, structured doctorates emerged in some subjects and professions, providing an additional influence on the PhD. The 'New Route' or 'Integrated' PhD (Park, 2005), for example, drew on new forms of doctorate in North America, was designed primarily for the international market, and contained a significant 'taught' element, balanced by a shorter thesis than was usual in the PhD. Much more significant than either the New Route or Integrated PhDs, however, were the professional doctorates. From their inception professional doctorates were highly structured and enabled candidates to undertake and apply research in their professional context.

Irrespective of subject, the structured PhD has a dual focus that develops both research and professional and personal² attributes, the latter being the focus of many employers of PhD graduates. One of its primary purposes was to address criticisms that some PhD graduates had not developed employment-related skills that prepared them for entry to non-academic jobs, a barrier to economic growth and global mobility.

Surprisingly, these significant changes and the ever-increasing emphasis on the employability of graduates have not been reflected in the PhD examination, which remains focused on research achievements and contrasts with other forms of assessment in UK higher education in its individuality and lack of openness.

2.4 Major changes to student assessment in UK higher education

The external peer review of academic practice introduced for all universities in the early 1990s (QAA, 2003) had a major impact on assessment. Criticisms that approaches to student assessment in all subjects were in need of improvement were in the public domain and universities were forced to take action to develop more transparent and objective assessment practices. These attempts to formalise, some might argue bureaucratise, assessment in UK higher education revolutionised strategies for student assessment in 'taught' programmes, i.e.

² Professional and personal attributes are grouped together hereafter, since they are so closely integrated.

Bachelor, Diploma and Masters degrees. Continuing priorities for universities and external agencies are to increase transparency for students on how their work is assessed and to provide them with timely feedback to improve their performance.

These events led to cohorts of students on taught programmes being assessed against intended learning outcomes (ILOs), sometimes described as criterion referencing, in a highly structured manner, a practice that has become widespread. Simultaneously there was a drive to increase the use of formative as well as summative assessment in taught programmes, partly as a result of raised student expectations. Students remain critical of the extent to which they receive timely and meaningful formative feedback that helps them to improve their work.

The evolution of the PhD has had a significant effect on the student experience and the purposes of the degree. Yet the PhD examination and the viva in particular (e.g. Morley, Leonard and David, 2002) have not changed to reflect either the evolution of the degree or the developments in assessment practice. The examination compares unfavourably with other forms of assessment, and is out of alignment with the now commonly accepted concepts of fairness, validity, reliability³ and consistency. The external peer reviews that have until recently been used to assure the quality of research degrees, have had a limited impact on the final examination, focusing mainly on the student experience and the role of supervisors and other support available for candidates. The only external guidance on the doctoral assessment process was included in the UK Quality Code (QAA, 2011), which outlined accepted good practice in the conduct of the current examination rather than attempting to change practice or suggest the examination should explicitly assess the range of graduate attributes contained in the UK-wide doctoral qualification descriptor (QAA, 2008). The descriptor represents the threshold standard to be achieved by any doctoral graduate and aims to encourage consistent doctoral outcomes in all disciplines, but by definition cannot include the subject-specific attributes sought by examiners, about which there seems to be a tacit agreement in most disciplines.

The obvious purpose of the PhD remains training the next generation of researchers, hence the universal requirement to make a contribution to knowledge in the relevant field. The

³ Reliability concerns the measure of agreement within and between different markers. Brown et al. say that specific marking schemes or criteria '*increase reliability*' (1997:234), but marking schemes do not exist for doctoral programmes and criteria are general rather than specific.

question remains whether the examination is fit for purpose and if the current format can be fully exploited by examiners to test both research and professional and personal attributes.

3 Rationale

The UK PhD examination currently consists of a two-part process: independent evaluation of the thesis (or equivalent, for example an artefact or composition with analytical commentary, hereafter `thesis') by two examiners followed by assessment of the candidate in the viva voce or oral examination (hereafter `viva'), almost always conducted in private. Examiners make their judgements on the basis of this process. The privacy of the examination reflects the individuality of the degree and its outcomes and has made it difficult to research this almost unique form of assessment. Criticisms of this arguably secretive examination focus both on its privacy and the potential inconsistency of examiners' judgements. The conduct of the examination is influenced not only by the candidate's work and viva performance but also by individual experiences and characteristics, the process itself and the potential outcomes.

Tinkler and Jackson (2004:8) confirm the centrality of the co-examiners' judgements in the examination:

We believe that examiners' perspectives about what constitutes a PhD are of central importance. Whilst examiners' views are sometimes informed by institutional policies and criteria and usually influenced by discipline knowledge and expectations, examining a PhD is still very much shaped by individual expectations and requirements.

This individuality raises questions about conformity with current expectations of assessment practice in higher education and has resulted in questions concerning the potential for inconsistent judgements. In the idiographic PhD examination it is more difficult to provide evidence of fairness, reliability and consistency or to demonstrate a shared understanding of the expectations that must be met to award a PhD. The viva in particular has been highly contested, seen as mysterious, overly private and an opportunity for examiners to engage in power-play, with one another and with the candidate (Wallace and Marsh, 2001). It is generally viewed as a phenomenon that has evaded the basic principles of valid and reliable assessment and `escaped' quality assurance practices (Morley et al., 2002; Tinkler and Jackson, 2002).

The first version of the QAA's code of practice for research degrees (QAA, 1999) included expectations that the examination *'should be clear and operated rigorously, fairly, reliably and consistently'* (p.12) and that all involved should share *'a clear understanding of the task'* (ibid).

Aiming to determine whether these expectations were being met, Denicolo, Fuller and Boulter (2000:62) found that academics were unsure how to define their expectations and that institutional assessment criteria were often vague, leaving room for interpretation. If widespread, these factors potentially undermined the concept of universally shared standards, and Denicolo et al. argued that in the absence of documentary evidence, it could not be assumed there was a consensus of standards among doctoral examiners and supervisors. Another consideration was the increasing numbers and diversity of doctoral graduates and the corresponding changes in demography among academic staff, who may have had '*little or no previous doctoral examining experience*' (p.63). Denicolo et al. emphasised that they were not seeking rigid rules for the doctoral examination, but to increase transparency, rigour and reliability as recommended by QAA.

When considered alongside arguments that the examination needs to change to reflect the PhD's changing status as a qualification for diverse employment destinations, the potential inconsistency of outcomes becomes doubly significant. The third dimension to this problem is that, given the PhD's enduring status as the highest academic qualification that confers a 'licence' to undertake research independently, apparent flaws in its assessment are potentially highly important for academic standards. However, as a degree undertaken 'by research' the PhD is different in kind from other degrees in higher education. This, together with the uniqueness of the process and of the thesis, may strengthen the case for its examination being different from other forms of assessment and explain the continued emphasis on evaluating research achievement.

The issues involved in the examination of the doctorate are both complex and controversial and for this reason '*the complexities of doctoral assessment have been relatively under-researched generally, both in Britain and internationally*' (Morley et al., 2002:268). This remains the case, particularly in relation to the viva, even though since 2000 there has been an increase in articles and books concerning the doctoral examination and some analysis of examiners' reports. The intricacies of the two-part examination, the relationship between the thesis and viva and their respective roles in identifying candidate attributes, remain relatively uninvestigated. The aim of this study is therefore to explore how the PhD examination enables examiners to judge if a candidate has met the requirements for the award of a PhD in their subject, taking account of the individual nature of the examination process and the candidate attributes, research-specific and professional and personal, sought by examiners. The research

questions reflect this objective and are closely linked to the three findings chapters (5, 6 and 7), which respectively explore: the examination process; the attributes sought by examiners; and the range of possible outcomes.

4 Research questions and study boundaries

4.1 Research questions

My research questions evolved with the title of the study and emerging data analysis. The thesis title reflects my overarching research question:

How does the PhD examination enable examiners to make judgements about the candidate and their work and is it still fit for purpose in the modern context?

The subsidiary research questions relate to my findings chapters:

i) *In the examination process, what role is played by the thesis, what are the purposes of the viva, and what is the relationship between them?* (Chapter 5)

ii) *What attributes do examiners seek in the thesis and the candidate and how are they demonstrated?* (Chapter 6)

iii) *How do different examination outcomes reflect candidates' achievements, and is a pass/fail judgement still appropriate?* (Chapter 7)

4.2 Boundaries of the study

Some of the boundaries for my project were created at the beginning to increase focus and manageability, whereas others emerged from implementing the research design. The three most important boundaries are explored below.

Focus on the PhD examination

Although the doctorate as a qualification has evolved and in the UK has two different forms – the PhD and the professional doctorate – this study focuses exclusively on the PhD for three main reasons: it has a long history of international currency and is relevant to all disciplines; professional doctorates exist in a limited number of fields; and many professional doctorates include interim assessments that contribute to the final outcome, therefore the process is not directly comparable to the PhD examination.

Limiting the study to the UK

While it would have been interesting to widen my research to countries outside the UK, any attempt to make comparisons with other PhD assessment systems was not feasible within the

scope of a doctoral project. However, the literature review considers international studies that compare the UK doctoral examination with other models, and any relevant comments from respondents regarding international examining have been included.

Access constraints: universities and subjects

Access constraints limited the range of universities and subjects I was able to sample and meant that opportunistic rather than systematic sampling occurred. The cases I eventually gained access to, however, provided rich data and some opportunity to compare subjects in two broad groups: Science, Technology, Engineering and Mathematics (STEM) and Arts, Humanities and Social Sciences (AHSS).

5 Thesis content and structure: summary

In **Chapter 2** I analyse the increasing regulation and external intervention in UK higher education and its contribution to the transformation of student assessment policy and practice. In reviewing the literature, I explore the reasons for the limited impact these changes have had on the PhD examination, despite the introduction of ground-breaking regulatory frameworks for doctoral education.

In **Chapter 3** I explore the examination in the context of the evolving doctorate and its changing purposes. Empirical research is principally limited to studies concerning: the purpose and conduct of the viva, including the effect of examiner behaviour on the candidate's experience; criteria used by examiners to evaluate the thesis; and the content and nature of examiners' reports, including evidence of both formative and summative assessment. Other publications question the examination's fitness for purpose in the modern context, commenting on the process, including power relations, and on maintaining standards of doctoral education in the UK and internationally. I explore the question of whether it is the individual, their work, or the training process that is being assessed, given the changing purposes of the doctorate. Finally, I review publications concerning the concept of 'doctorateness' and whether this elusive quality can be defined and assessed.

Chapter 4 outlines the rationale for my realist theoretical approach combined with case study methodology to generate data and discusses the appropriateness of my research design, based on ten bounded cases, for the task of exploring the PhD examination in depth. The

chapter introduces case studies of the ten viva observations, which were followed by semi-structured interviews with candidates, supervisors, examiners and independent chairs.

In **Chapter 5**, the first of the data chapters, I conduct an in-depth, holistic analysis of the examination process: the role played by the thesis, the purposes of the viva and the relationship between them. The analysis confirms a strong interdependency between the thesis and the viva as might be anticipated but adds new perspective to the two-stage process. The relative importance of the thesis and viva in contributing to the final judgement are explored and confirm the individuality of the examination.

Chapter 6 explores the heterogeneous attributes sought by examiners in the thesis and the candidate. One of the more striking findings in this chapter was the extent of agreement among examiners about the qualities sought in candidates and their work, even though disciplinary considerations were apparent when a candidate's contribution to the field was evaluated. However, in making higher-level judgements, such as whether overall a candidate had 'done enough' to merit the award of a PhD, there was considerable alignment.

Chapter 7 focuses on three elements of examination outcomes: diversity, comparability and grading. Here I explore the extent to which the diversity of candidates and their attainments has the potential to undermine consistent standards in the PhD and how far examination outcomes and graduates can be considered comparable. I explore the question of whether the PhD should be graded, with some contradictory results.

Chapter 8 considers the most striking findings from the study in the context of the literature. It discusses the key factors in the study, demonstrating the interdependency of different elements of the examination in arriving at the final result. It illustrates the complexity inherent in examiners' judgements and the range of influences affecting each examination, including expectations in the field, consistency and maintaining standards. Here I also analyse the effectiveness of the research design and reflect on the challenges and limitations of the study

In the thesis conclusion, **Chapter 9**, I explore to what extent the research question has been answered, and suggest how the results of this study might be used to influence policy and practice in the UK. Finally, I suggest possibilities for future research.

Chapter 2: The impact of policy changes on assessment in higher education

The *raison d'être* of a higher education is that it provides a foundation on which a lifetime of learning in work and other social settings can be built. Whatever else it achieves, it must equip students to learn beyond the academy once the infrastructure of teachers, courses and formal assessment is no longer available. This is a formidable challenge and it competes with a number of other goals to receive the attention it deserves in a university education. (Boud and Falchikov, 2006:2)

1 Introduction

The early 1990s saw one of the most fundamental changes ever to occur in UK higher education. The 1992 Further and Higher Education Act removed the differentiation between polytechnics and universities (the so-called 'binary divide'), blurring the boundaries between polytechnics that specialised in 'applied' degrees and universities offering more theoretical and research-led education. Giving all higher education institutions university status created a new dichotomy of 'pre-1992' and 'post-1992' universities, increased competition for doctoral students and contributed to the growth in doctoral numbers described in Chapter 1. This unprecedented external intervention in UK universities contributed significantly to the transformation of student assessment.

The 1992 Act gave the UK HE funding bodies statutory responsibility for assuring the quality of all UK degrees and brought about seismic changes that continue to affect the autonomy of all UK higher education institutions. These included the introduction of UK-wide 'quality assurance' mechanisms that aimed to regulate the standards of all degrees through external peer review. Described by Wagner (1993:280) as a '*major shock*' to universities, the merger of existing processes, including self-regulation, into a 'blended' quality assurance system suitable for the new non-binary environment was complex and controversial (Watson, 1995). Establishment of the QAA, under whose stewardship a UK-wide framework of guidance, an 'Academic Infrastructure' (AI) was created, had far-reaching consequences and led to new ways of thinking about teaching, learning and assessment (Griffiths, 2014).

This chapter demonstrates the impact on student assessment of implementing major policy changes, including the introduction of pedagogic terminology and more technical approaches to assessing students. It also considers why the new assessment regime has not permeated doctoral education. Whereas assessment in universities had previously focused predominantly

on curriculum coverage, the introduction of intended student learning outcomes (ILOs) added a completely new dimension that coincided with the emphasis on 'transferable' skills for all graduates. The growing importance of ILOs and transferable skills was evident in a ground-breaking doctoral qualification descriptor combining research and professional attributes capable of being tested in the examination. Finally, the notion of PhD standards and the feasibility of subject-specific ILOs are explored in the context of a particular discipline. I conclude that advances in assessment, which achieved real impact in taught programmes, have affected the PhD only marginally.

2 Policy changes designed to influence academic practice

The QAA's responsibilities for external peer review of teaching, learning and assessment included development of a Code of Practice in partnership with universities. The introduction of this ground-breaking framework for academic practice, which included sections on student assessment, external examining and research degrees, was highly significant for university autonomy. One of the AI's most influential components was a Framework for Higher Education Qualifications (QAA, 2001, 2nd edition 2008). The FHEQ aimed to set standards for all degrees, including doctorates, and included qualification descriptors containing learning outcomes encompassing academic attributes and 'transferable' skills. Universities were expected to adhere to these national descriptors and to use them internally and in external examining. Creation of a doctoral qualification descriptor was a major departure that for the first time set out learning outcomes for doctoral degrees.

Driven by the new quality assurance regime; the foregrounding of student learning objectives; the impact of assessment on student learning; and the influences of '*the employability and graduate skills agenda*' (Bloxham and Boyd, 2007:4), UK universities introduced more detailed and systematic internal policies concerning student assessment (Bryan and Clegg, 2006). They aimed to increase transparency, validity and reliability and to demonstrate that students had acquired the academic knowledge and transferable skills enshrined in the qualification descriptors. As Bloxham and Boyd (2007:3) later observed '*The contemporary environment of higher education means that assessment cannot carry on unaltered; it is subject to too many pressures and influences which create a force for change*'.

Universities largely endorsed many of the AI reference points introduced in the early 2000s, including those for research degrees. These UK-wide 'benchmarks' aimed to encourage

effective and consistent practices. They contributed to standard-setting and focused on the interdependent relationship between teaching, learning and assessment.

3 New ways of thinking about teaching, learning and assessment

As research into higher education in the UK and internationally grew, the conceptualisation and conduct of assessment in taught programmes changed considerably. This had little impact on the PhD and its examination, however, which *'is notoriously difficult to research'* (Morley 2004:91).

Academic research into teaching, learning and assessment and the relationship between them has a long history in secondary education and remains prolific. Murphy suggests that in school education *'the reform of assessment systems has been seen as a key factor in bringing about improvements in student learning'...*[while]...*'much of higher education has tended to plod along'* (2006:38). Nevertheless, changes in pedagogical thinking in universities resulted in *'many of the trends around assessment...[becoming] common to both schooling and higher education, including an intensified and high stakes assessment regime, and concerns about standards and equity'* (Leathwood 2005:308). The many challenges faced by university teachers included *'the professionalisation of academic staff in relation to teaching, learning and assessment'* (Bloxham and Boyd, 2007:5). Academics are now required to possess theoretical knowledge commensurate to their practical experience and to engage in continuing professional development (CPD). Supported by a canon of research, a sea change has occurred in how assessment is understood, and in the underlying principles involved, including fairness, reliability, validity and other concepts. The *'primacy'* (Bloxham and Boyd, 2007:3) of assessment in the teaching/learning/assessment relationship is now well established. Important shifts in emphasis have embedded the understanding that, *'to a large extent, assessment activity in higher education is the learning activity'* (Brown, with Bull and Pendlebury, 1997:7), that assessment is pivotal to the way students learn and needs to test diverse cognitive skills.

The significant body of academic literature concerning teaching, learning and assessment in higher education is both theoretical and practical. Rarely is assessment explored in isolation. The literature acknowledges the interdependency of the three practices (Boud and Falchikov, 2006, Biggs and Tang, 2007) and explores the wide-ranging assessment methods employed in UK higher education, analysing why different forms are appropriate depending on the student

attributes sought, the subject context and the type of learning intended (Brown, 1997; Biggs, 2003). Dissemination of emerging research was strongly motivated by national quality assurance initiatives such as Subject Review (QAA, 2003).

The changed culture in pedagogical understanding and research in educational assessment brought about criterion referencing and anonymous marking in taught programmes to help demonstrate fair, reliable and valid assessment, seeking to replace the 'norm' referencing, which had previously been relatively uncontroversial⁴. The new priority became to assess students against specific objectives:

Norm-referenced interpretations of assessment outcomes can be of benefit in making selection decisions, but in educational assessments, particularly those intended to support learning, it is far less important to know where an individual is in some rank order than it is to find out what they have learnt. (William, 2004:356)

By contrast, the final, summative PhD examination, designed to assess individual learning principally driven by the demands of the research environment, was not subject to such detailed scrutiny, nor was it affected by technical or theoretical approaches. According to Delamont, Parr and Atkinson (1998), this was in part due to the practical difficulties of researching the examination, including the barriers to direct observation of the viva and the confidentiality of examiners' reports.

In taught programmes, interventionist approaches to assessment continued. Terminology was introduced to differentiate between forms and purposes of assessment. Methods were classified as diagnostic, formative or summative, although the categories were acknowledged not to be mutually exclusive. These terms became familiar at subject level (Brown, 1997; Rust, 2002; Biggs and Tang, 2007). A more analytical approach to the purposes of assessment and its impact on student learning led to greater creativity and innovation. Bloxham and Boyd (2007:7) emphasise the continuing value of professional and disciplinary judgement in assessment, a concept equally applicable to taught and research degrees that has particular relevance for judging the outcomes of the highly individual PhD examination. They foreground the critical role of teachers in matching assessment to context:

⁴ 'Criterion' referencing infers that students in a cohort are assessed against pre-determined criteria that reflect what they should have learned. 'Norm' referencing, on the other hand, where *'performance against some standard has been interpreted...by comparing the performance of an individual...against [another] group of individuals'* (William, 2004:356) assumes that the performance of different student cohorts in a given assessment task does not differ greatly year on year and ranks students in order of their scores within the overall cohort: *'a "normal distribution" in statistical terminology'* (Australian Universities Teaching Committee, 2002).

Although the growing evidence base of research on assessment provides a useful basis on which to build and review practice, it leaves the onus on tutors and teaching teams to develop and critically evaluate assessment processes and procedures...within their local context...While the evidence base can inform institutional and departmental policy, it will require mediation to suit local contexts and student groups.

Emerging literature on doctoral assessment was influenced by policy-makers' assumptions that assessment based on learning objectives could be extended to doctoral degrees. These expectations had been influenced partly by the merger of cultures brought about by the 1992 Act and the introduction of the external quality assurance measures already described, themselves a contributory factor in the revolution in teaching, learning and assessment and recognition of the relationship between them. Quality assurance mechanisms aimed to introduce more consistent standards and accelerated the implementation of new assessment practices.

4 The impact of the doctoral qualification descriptor

The introduction of the doctoral qualification descriptor as part of the original FHEQ was arguably the greatest intrusion into universities' academic autonomy over doctoral degree standards, encapsulating UK-wide, absolute, threshold standards for all doctorates. Despite its significance, the descriptor was not contested, perhaps because it was perceived to fulfil a long-standing need to establish some generic characteristics of doctoral graduates, or because supervisors and examiners were not obliged to use it and may have been unaware of it.

As Tinkler and Jackson (2000:168) observe: *'policy and practice are not synonymous; what is set out in institutional policy is not necessarily adhered to in practice nor is it, in some instances, even widely consulted by academic staff'*. The doctoral descriptor was comprehensive, yet general enough to apply to all doctoral graduates. In 2008, when QAA was updating the FHEQ and consulted universities on potential changes to the descriptor, respondents identified few problems with the core attributes, only suggesting additional text (Table 2.1, 4.18.1 to 4.18.6) to augment the existing content and make it more inclusive of all doctorates:

4.18 Doctoral degrees are awarded to students who have demonstrated:

- *the creation and interpretation of new knowledge, through original research or other advanced scholarship, of a quality to satisfy peer review, extend the forefront of the discipline, and merit publication*
- *a systematic acquisition and understanding of a substantial body of knowledge which is at the forefront of an academic discipline or area of professional practice*
- *the general ability to conceptualise, design and implement a project for the generation of new knowledge, applications or understanding at the forefront of the discipline, and to adjust the project design in the light of unforeseen problems*
- *a detailed understanding of applicable techniques for research and advanced academic enquiry.*

Typically, holders of the qualification will be able to:

- *make informed judgements on complex issues in specialist fields, often in the absence of complete data, and be able to communicate their ideas and conclusions clearly and effectively to specialist and non-specialist audiences*
- *continue to undertake pure and/or applied research and development at an advanced level, contributing substantially to the development of new techniques, ideas or approaches.*

And holders will have:

- *the qualities and transferable skills necessary for employment requiring the exercise of personal responsibility and largely autonomous initiative in complex and unpredictable situations, in professional or equivalent environments.*

4.18.1 Doctoral degrees are awarded for the creation and interpretation, construction and/ or exposition of knowledge which extends the forefront of a discipline, usually through original research.

4.18.2 Holders of doctoral degrees are able to conceptualise, design and implement projects for the generation of significant new knowledge and/or understanding. Holders of doctoral degrees have the qualities needed for employment that require both the ability to make informed judgements on complex issues in specialist fields and an innovative approach to tackling and solving problems.

4.18.3 Doctoral programmes that may have a substantial taught element in addition to the research component (for example, professional doctorates), lead usually to awards which include the name of the discipline in their title (for example, EdD for Doctor of Education or DClInPsy for Doctor of Clinical Psychology). Professional doctorates aim to develop an individual's professional practice and to support them in producing a contribution to (professional) knowledge.

4.18.4 The titles PhD and DPhil are commonly used for doctoral degrees awarded on the basis of original research.

4.18.5 Achievement of outcomes consistent with the qualification descriptor for the doctoral degree normally requires study equivalent to three full-time calendar years.

4.18.6 Higher doctorates may be awarded in recognition of a substantial body of original research undertaken over the course of many years. Typically a portfolio of work that has been previously published in a peer-refereed context is submitted for assessment. Most degree awarding bodies restrict candidacy to graduates or their own academic staff of several years' standing.

Table 2.1 Doctoral qualification descriptor (QAA, 2008)

The descriptor proved a key resource for augmenting university regulations for doctoral assessment. The attributes it contained were sufficient to convey the threshold standards conceived of by universities, as well as stakeholders such as the funding and research councils.

It also had a major influence on the European doctoral descriptor (Bologna Working Group on Qualifications Frameworks, 2005), which is almost identical.

As Denicolo points out (2003:86) *'The criteria for attainment of a doctorate are most commonly couched at a very general level'*. While the descriptor enshrined typical attributes for doctoral graduates that could be referenced during the examination process, potentially increasing consistency of judgements, it did not solve the problem of agreement among disciplines about outcomes or differing examiner expectations. In exercising their professional judgement, PhD examiners are potentially open to criticisms of subjectivity, although the arguments are complex. Delamont et al. (2000:178) suggest it is not possible to apply a learning objectives approach to the PhD because of the *'indeterminate assumptions'* that exist about quality, especially in the context of disciplinary differences, whereas others address more fundamental issues concerning the doctorate itself. Even since publication of the descriptor, Denicolo (2003) suggests that until there is a shared understanding about what possession of a doctoral degree represents, taking account of its multiple forms and purposes, its assessment will remain unsatisfactory.

Crucially, the UK descriptor aligns with international frameworks, including the Australian Qualifications Framework (AQF, 2013). This alignment demonstrates international parity of expectations and achievement for doctoral graduates, of prime importance for the UK in the context of international competition for students (Schwab, 2014; Kemp et al., 2008). Learning outcomes for doctoral degrees have been embraced in the United States, featuring in the evaluation of doctoral education (Maki, 2006; Funk and Klomparens, 2006) and in progress examinations *'to assess how well doctoral students integrate, transfer and apply their learning'* (Maki, 2006: 6).

The AI, especially the Code of Practice and the FHEQ, played a significant role in the external review of all degrees. Deficiencies in assessment practice, despite developments in policy and practice, continued to be identified by reviewers, especially in taught degrees.

5 Peer review of doctoral assessment

External peer review of assessment of doctoral candidates in the first decade of the 21st century revealed some concerning outcomes, showing lack of alignment with current assessment practice in higher education.

Aiming to assess how the revised Code of Practice for research degrees (QAA, 2004) was being implemented in universities, during 2005-06 QAA launched a desk-based, special peer review. While the results of this process were overall highly positive, student assessment was by far the weakest area of provision. This was unexpected and of some concern. The extent of the problem was highlighted in the England and Northern Ireland report's conclusions, which suggested that *'finding ways of assuring fairness and consistency in the oral examination'* and *'implementing more detailed assessment criteria'* were areas that institutions found *'challenging'* (QAA, 2007:23).

In 2012 QAA published another, broader-based report on postgraduate research degrees using data from visits to around 60 diverse institutions between 2009 and 2011, rather than relying on institutions' self-evaluations. On assessment, the authors first describe practices in appointing and preparing examiners without venturing to make judgements about effective or poor practice. However, several audit teams concur with the 2007 report by identifying the positive impact of independent chairs in vivas as *'a feature of good practice which provides for equity of treatment and robustness of outcome'* (QAA, 2012:27). In one university that had not implemented such a system, auditors suggested *'it would be desirable to consider whether there might be an advantage in use of independent chairs'* (p.30). In another report the auditors *'observed that the use of independent chairs in only some examinations had the potential to promote inconsistency'*, while the report on a third university suggests:

There should be a review of the use of independent chairs to further secure fairness and consistency in postgraduate research examination[s]...based upon concerns that current practice was for the independent chair to withdraw while the examiners were reaching their judgement. (QAA, 2012:30)

Reviewers' comments suggest that the use of independent chairs can increase fairness and consistency by ensuring that examiners follow the university's rules for conduct of the viva (Poole, 2015). However, they also warn that inconsistencies could remain if independent chairs are not present in all cases, or where their role is unclear.

In both 2007 and 2012, the references by QAA reviewers to the desirability of fairness and consistency, particularly use of independent chairs, supported continuing criticisms of the PhD examination in the academic arena. Nevertheless, universities were relatively successful in demonstrating that their doctoral examination practices aligned with their regulations and guidelines and that these reflected those enshrined in editions of the UK Quality Code for

research degrees (QAA, 2004; 2011, section B11), which provides general expectations for the conduct of all doctoral degrees.

6 The student voice

The authority of the student voice has been growing steadily since quality assurance mechanisms were first introduced and student views of their experiences now have considerable influence. There is an expectation that all students, including doctoral candidates, will be regularly surveyed on their learning experience, including assessment, which continues, marginally, to be one of the least satisfactory aspects of the student experience in taught programmes. This has been demonstrated through annual surveys of student opinion, such as the National Student Survey (NSS – for undergraduates) and Postgraduate Taught Experience Survey (PTES). The evaluation of assessment practices is a complex process and student views constitute one of several measures, yet NSS and PTES results continue to suggest students are less satisfied with assessment than with other elements of their experience. The consistently weaker results for assessment than for other areas suggest that formalisation of assessment policy nationally and at university levels may not yet have become completely embedded.

While research students are more positive about assessment, they are also happier with some other elements of their experience, as evidenced since published results have been available (biennial since 2009). The Postgraduate Research Experience Survey (PRES) response rates are relatively low, although have improved steadily from 29% in 2009 to 41% in 2015. The small proportion of respondents in each year who have experienced their final assessment at the time of the survey is a limitation. However, the results from 2009 and thereafter suggest that in the UK candidates are generally satisfied with their experience of progress tests and final assessment. In a commentary on the numerical report of PRES (Park, 2009:4), comparisons were made with the outcomes of a similar survey in Australia, showing UK students were less satisfied than their Australian counterparts concerning 'goals and standards' but almost as positive about the final examination, *'research students who were eligible to answer the questions in PRES on thesis examination had generally very positive views, particularly on the fairness and timescale of the...process'*.

In 2009, compared with other scales (mean scale scores: 1= always definitely disagree to 5 = always definitely agree), questions on progress and assessment (labelled 'goals and standards') were answered positively by research students, with an average of 3.8 compared with

supervision (4.0), skills development (4.0), infrastructure (3.65), intellectual climate (3.4) and professional development and career (3.15). Most students who had experienced the thesis examination answered the relevant question positively, as summarised in Table 2.2:

	Disagree	Neutral	Agree	N
6.a.i The thesis examination process was fair	12.7%	9.4%	77.9%	716
6.a.ii The examination of my thesis was completed in a reasonable time scale	14.3%	9.3%	76.4%	720
6.a.iii I was given adequate support and guidance in preparation for my viva voce	22.4%	16.0%	61.6%	718
6.a.iv I was given adequate support and guidance to make any changes to my thesis following my viva voce	19.4%	13.7%	66.9%	665

Table 2.2: 2009 PRES report (HEA, 2009; Table 11: Thesis Examination 2009)

The area showing the lowest levels of satisfaction was preparation for the viva, with more than 20% of those surveyed feeling they had not received sufficient support and guidance beforehand. Results were more positive in 2011, when 4.6% of the respondents (n=1,401) answered questions in this scale. For example, 84% of respondents said they agreed the examination process was fair, compared with 78% in 2009. On support for preparing for the examination, positive responses were lower, at 71% compared with 62% in 2009. The greatest increase in positivity occurred in responses to the question about support for post-viva thesis corrections, with a 10% increase from 67% in 2009 to 77% in 2011 (Hodson and Buckley, 2011).

In 2013 the survey was re-designed so results cannot be compared. Candidates' responses to assessment experiences remain positive, however, with *'those who understand the required standard for their thesis increasing from 75.5% of first years to 83.5% of fourth years'* (Bennet and Turner, 2013:29). Clarity about *'final assessment procedures'* also increased, *'from 70.9% in year 1 to 79.4% in year 4'* (ibid). Using the mean scales scores, in 2015, survey data showed a mean score of 4.07 for progress and assessment, compared with 4.03 in 2013 and that the mean percentage agreement with questions concerning progress and assessment was 79%, compared with 78% in 2013.

Numbers of PhD candidates who have experienced the final examination when the survey occurs are small but the results are interesting. They suggest that doctoral students are rather more satisfied with their assessment experience than their counterparts on *'taught'* programmes and that universities have responded to survey outcomes by working to improve assessment procedures. When students were asked directly about the fairness of the examination process in 2011, 84% agreed it was fair. It would be interesting to know if there

was any correlation between examination outcomes and graduates' opinions concerning fairness, i.e. if those who passed with no or minor corrections were more likely to agree the process was fair than those who were asked to complete major corrections or were awarded a lower degree. These data are not currently available.

The contrast between the opinion of students on taught programmes and research degrees concerning their assessment experiences is striking. In taught programmes, while improvements are demonstrable, there is continuing criticism of some elements of assessment, especially the amount and content of formative feedback. On the other hand, research students appear more satisfied overall with their assessment experiences than the literature would suggest, although both the individual nature of the PhD and the context of doctoral training as an induction into the academic community may influence their view. The introduction of external quality assurance processes had a major impact on universities' assessment policies, and in some cases, practices. They increased the significance of intended student learning outcomes (ILOs), both subject-related and personal, as external peer reviewers scrutinised what was expected of students and the effectiveness of how those attributes were assessed. The next section explores the increasing importance of ILOs in the doctorate.

7 Intended student learning outcomes (ILOs) in the doctorate

Statements of ILOs, or 'learning objectives' were a major feature in external reviews of universities, and simultaneously became fundamental to the widespread use of criterion-referenced assessment. Originally intended to enhance programme design, ILOs enabled teachers to articulate attributes they expected students to acquire during a module or programme, whether subject-related or transferable to other contexts. Over time, ILOs encompassed multiple expectations – academic, government- and employer-related and student-focused – especially employment-related transferable skills.

QAA subject benchmarks for taught degrees created by the academic community in the relevant field, or degree specifications set by professional bodies or learned societies, describe the ILOs expected of graduates in their discipline, including subject knowledge, transferable skills, progression levels and the attributes required for any form of professional accreditation. These also act as a proxy for degree 'standards'.

The attributes in the doctoral qualification descriptor (Table 2.1) have frequently been used to shape institutional assessment guidance and regulations, but it is not known to what extent doctoral examiners refer to them directly, either in their original form or when embedded in institutional guidance. The attributes describe both research and professional characteristics that are intended to be recognisable by the academic community in any subject (Lovitts, 2007) and to encompass the higher education sector's expectations of what any doctoral graduate is capable of upon graduation. As such, they assume the role of doctoral ILOs. Kiley (2009:33), for example, supports a strong relationship between stated doctoral learning outcomes and assessment of candidates, suggesting that:

The final assessment processes for doctoral education should be aligned with the proposed learning outcomes and should...contribute to the learner's development and understanding of being a researcher.

On the other hand, she also acknowledges that *'to articulate the learning outcomes of the doctorate adequately is far from easy, even at the broadest level'*.

The question of how candidate attributes are evaluated in the doctoral examination is difficult to address for two principal reasons. First, the ILOs cannot always be directly linked to examiners' judgements, due to the individuality of the examination and its unique circumstances as discussed in Chapter 1. Second, the lack of specific learning outcomes rooted in the field of study (such as the doctoral equivalent of a subject benchmark in an undergraduate degree) relies on the professional judgement and interpretation of individual examiners (Bloxham and Boyd, 2007). It is possible that, due to the unique circumstances of the PhD where the focus throughout is on individuality, independent learning and enquiry, more explicit, subject-related ILOs at doctoral level might not be feasible. Nevertheless, at least one subject community has attempted to create them.

Critics of criterion-referenced assessment using ILOs argue against narrow learning outcomes at any level. Leathwood (2005:320), for example, suggests that outcomes intended to *'foster intellectual independence'* are sometimes replaced by *'more prescriptive and specific outcomes'* (p.310). In the PhD, Delamont et al. (2000:178) make a compelling argument for retaining implicitness in determining *'what "counts" as research and...quality or originality'*, suggesting that when examiners describe what they are seeking in candidates they *'do not reproduce the kind of checklist with which universities supply external examiners'* or seek to use formal learning *'"objectives" or outcomes'*. This interpretation of examining approaches

has authenticity in a practical context, where examiners are judging something unique in the case of both the work and the individual.

Politically, a key feature of introducing ILOs was the opportunity to incorporate transferable skills in expectations of graduate outcomes previously confined to academic, or '*cognitive*' achievements (NCIHE, 1997:34). During the 1990s, governments had begun to take a greater interest in the tangible economic and sociological benefits of the continuing public investment in higher education and to raise their ambitions for participation (Morley, 2004).

Simultaneously, the combination of external scrutiny and internal quality assurance systems increased the accountability of universities for the public funding they received. With rapidly growing student numbers, the government had turned its attention to the '*transferable skills*' expected of all graduates (NCIHE, 1997).

8 The challenge of defining and assessing '*transferable*' skills

The governmental interest in universities became more intrusive as the quality assurance regime intensified. As a result, Sir Ron Dearing was commissioned by the UK governments to chair a committee of enquiry

To make recommendations on how the purposes, shape, structure, size and funding of higher education, including support for students, should develop to meet the needs of the United Kingdom over the next 20 years, recognising that higher education embraces teaching, learning, scholarship and research. (NCIHE, 1997:1)

The committee's seminal report '*Higher Education in the Learning Society*', known informally as the Dearing report (NCIHE, 1997) had far-reaching implications for universities, particularly the requirement: '*learning should be increasingly responsive to employment needs and include the development of general skills, widely valued in employment*' (ibid). This built directly on the Enterprise in Higher Education programme, which '*aimed to...increase the effectiveness of Higher Education (HE) in preparing students for working life*' (Burniston, Rodger and Brass, 1999:1), a government ambition to elicit cultural change that would increase graduate '*employability*'. Dearing identified skills that would enable graduates to thrive in the future labour market and coined expressions that have become ubiquitous: '*portfolio careers*', '*lifelong learning*' and '*transferable skills*' (NCIHE, 1997:56;57;34).

Setting and assessing subject-specific ILOs in universities was complicated by this political emphasis on transferable skills. Dearing notes the lack of '*commonality*' among employer

views regarding *'personal transferable skills'* (NCIHE, 1997:34), which still pertains.

Unsurprisingly, given the amount of diversity among graduate employers, the Committee was unable to establish a *'consensus'* or *'definitive list'* of skills sought, or agreement about what graduate skills were lacking (p.133). This is a rare reference in a policy report to the very diverse requirements of graduate employers. Dearing did, however, identify some professional and personal attributes valued by many, including critical thinking ability, and communication and technology skills, and suggested that *'The largest single expression of dissatisfaction comes from the 25 per cent [of employers] who would like graduates to have better communication skills'* (p.34). Conversely, most students felt *'their skills [had] improved during their time in higher education'* (ibid).

A major innovation for doctoral degrees occurred when in 2001 the UK Research Councils developed their Joint Skills Statement (JSS), later subsumed in the Research Degrees Code of Practice (QAA, 2004). The ground-breaking JSS emerged from agreement among the Research Councils about the generic skills all doctoral candidates should acquire. Subsequently, SET for Success (Roberts, 2002), identified a deficit in the transferable skills of many STEM PhD graduates. The impact of the two publications was that targeted additional public funding for structured training in all subjects aimed at developing transferable skills was bestowed on universities with relatively large numbers of research students. The Roberts report in particular had a major and continuing effect on how doctoral training in all UK universities is conceived and implemented, although initially it was divisive as the funding mainly benefitted research-intensive institutions.

While the *'Roberts'* funding provided additional funding and increased awareness of the need for structured doctoral training, it could not be used for research skills training i.e. attributes in the *'Research skills and techniques'* or *'Research environment'* categories (QAA, 2004:34). Arguably this inadvertently created a false dichotomy by suggesting that the attributes gained through learning about and doing research were not transferable to other contexts.

Qualities such as *'critical thinking'*, *'ethical practice'*, *'creativity'*, and *'independent problem solving'* among many others (Phillips and Pugh, 2010), have become virtually generic expectations of all graduates, although the doctoral qualification descriptor (Table 2.1), attempted to differentiate the research and higher level attributes that could be expected of doctoral graduates. Lovitts (2007) makes a case for a learning outcomes approach in doctoral

education, arguing that making explicit the expectations and performance involved in doctoral success is likely to improve candidates' engagement in *'intellectual and skill development'*, as well as making them *'better able to self-assess and correct deficiencies, and...to demonstrate what they know and can do'* (p.24). She also suggests that clarifying standards and performance expectations in the context of the discipline increases fairness and accountability.

These arguments remain salient, especially for the PhD, where graduates are expected to possess sophisticated transferable skills commensurate with their high level academic qualification. Most of the literature argues transferable skills are best acquired through academic study (Biggs and Tang, 2007). This can be extrapolated to doctoral candidates, who develop some of the most sought-after attributes by learning about and doing research in a complex environment while undertaking cohort-based training to gain specialist or generalist knowledge. The expectation that new PhD graduates are equipped with a range of highly developed transferable skills has contributed to the current emphasis on skills training. Criticisms of PhD graduates' preparedness for employment remain a constant refrain in public discourse (HEC, 2012), however. Perhaps this is inevitable given the individuality of the degree, the uniqueness of each candidate's experience and, as Dearing identified, the particularly diverse needs of employers of PhDs, who range from SMEs to large international corporations.

Each PhD graduate has had a unique experience and development path and has acquired a different set of attributes, influenced by their field of study. This emphasises the wisdom of retaining a doctoral qualification descriptor that sets threshold standards. Universities' assessment criteria for the final examination often reflect the research and professional attributes set out in the doctoral qualification descriptor that acts as a proxy for doctoral ILOs, yet in practice the examination does not explicitly test transferable skills. PhD examiners are well placed to evaluate multiple attributes during the two-part final examination and while their expectations of candidates are often implicit (Lovitts, 2007), in practice professional and personal attributes are tested alongside research achievements.

9 PhD standards and subject communities

Unlike most learned societies, the International Union of Biochemistry and Molecular Biology has developed *'Standards for Doctoral Degrees in the Molecular Biosciences'* (IUBMB, 2011)⁵.

⁵ Contemporary with the current study, only two other professional associations appear to have attempted a similar exercise, as follows. The Association of Medical Schools in Europe and the World

Recognising that formalising subject-specific ILOs would contribute substantially to the validity of assessment and to maintaining standards, the IUBMB initiated a discussion among their academic community. After international consultation, the Standards were first published in 1989 and revised in 2000. The feasibility of introducing a learning outcomes approach for the PhD has rarely been addressed, perhaps understandably given the individuality of each degree and the unique circumstances surrounding it. The IUBMB's attempt to introduce such an approach, however, contributes to fulfilling expectations of fairness, validity, reliability and consistency, extending the use of criterion referencing to the PhD (Shuell, 1986; Biggs and Tang, 1997).

The Standards aimed to define *'the qualities and skills expected of doctoral graduates by the international community of molecular bioscientists'* (Wood and Vella, 2000:771), indicating the global reach of the PhD and the IUBMB's aim to create a common understanding of the qualities sought in graduates in the field. Identifying the rationale for revising the Standards, Wood and Vella note the PhD's primary role in training candidates to conduct independent research. While the Standards prioritise research attributes and outputs, they also recognise the need for candidates to acquire *'transferable knowledge and skills'* particularly in relation to non-academic employment (IUBMB, 2011:4).

The IUBMB considered the PhD Standards necessary because *'the standards to be achieved ...have been more intuitive than objective in nature, and have led to great variation in the abilities and skills acquired by graduates'* (Wood and Vella, 2000:771). Observing that PhDs in biochemistry and life sciences are conferred by many and diverse institutions and that it is commonly assumed that academic standards are universal, the IUBMB argue that while external assessment is fundamental to the PhD and its examination, the current process *'does not constitute a clearly defined international standard for comparison of graduates'* (ibid). These are crucial observations in the context of global growth in PhD graduates and can be extrapolated to other subjects, raising questions about why other academic communities have not identified the need to create disciplinary standards rather than relying on intuition or assumptions than standards are similar irrespective of context.

Wood and Vella (2000:771) do not suggest standards are unsatisfactory, as is evident from the

Federation for Medical Education have jointly developed proposed Standards for PhD education in biomedicine and health sciences in Europe (Orpheus, 2012), and the most recent British Psychological Society guidelines contain advice concerning the examination process which include a brief summary of attributes to be attained by PhD graduates (BPS, 2017).

international and wide-ranging career destinations of biosciences PhDs. They acknowledge, however, that undesirable outcomes may arise from conflicting expectations among examiners and that such concerns led to the IUBMB's initiative. The Foreword to the Standards (IUBMB, 2011:3) emphasises the complexities of assessing highly interdisciplinary life sciences degrees and asserts that as a result of the changes and growth affecting molecular biosciences in the 20th and 21st centuries:

The thousands of scientists qualified to supervise professional training in the Molecular Biosciences now comprise a heterogeneous group...the informal methods of the past do not permit similar [doctoral] standards among nations, within a country, or even within the same institution.

This statement reinforces concerns about inconsistency among arbiters of standards and emphasizes the potential for extrapolating the IUBMB's approach to other fields.

While acknowledging that '*progress in scientific knowledge and understanding does not come equally from all members of the profession*' (ibid), the IUBMB nevertheless suggests that society assumes equivalent standards among universities, a view not universally shared (House of Commons Innovation, Universities, Science and Skills Committee, 2009). The IUBMB accepts some inequalities of knowledge and understanding among individuals but does not suggest this arises from or leads to significant differences in standards of PhD graduates, or of universities. The IUBMB aims the Standards should:

Emphasize abilities...characteristic of those awarded a doctor's degree in a Molecular Bioscience, suggest how these abilities may be acquired...how their attainment may be assessed, and propose criteria for the overall evaluation of candidates (IUBMB, 2011:4)

This clarifies the purposes of the statement and the IUBMB's intention that the Standards should provide practical rather than theoretical guidance. As such, they aim to support universities and other higher education sector organisations in maintaining standards for research degrees and to help examiners and candidates to know what is expected of them. In the context of this study, these appear to be desirable objectives.

The IUBMB is clear that the fundamental condition for the award of a PhD remains that the candidate should produce original work, or '*constitute a significant contribution to advancement of basic or applied knowledge*' (IUBMB, 2011:5). This is addressed in a general section that precedes the Standards themselves and which makes clear the expectations of doctoral research are beyond those of Bachelors or Masters degrees.

The six Standards set out in the document are additional to this underlying requirement and summarise the detailed attributes to be met by doctoral graduates in this broad-ranging field. The Standards are supplemented by sections concerning: integrity in science, the role of formal training, the responsibility of the research supervisor, the responsibility of other academics, the responsibility of the candidate, funding of research training, duration of training and the thesis, each designed to provide guidance for all those involved. A striking feature of the Standards is that they effortlessly merge research and transferable skills, giving the impression of a rounded graduate with a portfolio of employment-relevant skills, irrespective of chosen career.

The 'Standards' emerged from a global consultation in the IUBMB community, giving them international standing. Examiners may use these standards against which to benchmark candidates if they wish, but it is unknown whether the majority of the molecular biosciences community are aware of the Standards, or if they are widely used by those supporting, training, supervising and examining PhD candidates in the field. If the Standards have been embraced by the subject community, they may have increased consistency and reliability, and possibly validity; provided candidates with a deeper understanding of what is expected of them in the thesis and viva; become an international benchmark for PhD examiners.

10 Conclusion

The transformation of student assessment in UK higher education aimed to increase consistency, fairness and understanding of its pedagogical importance. The introduction of explicit ILOs, shared with students (QAA, 2003) and used in assessment, was particularly revolutionary. Implementation of change has been possible for several reasons, including: adoption of the fundamental principles of assessment and communicating assessment tasks to students; the widespread introduction of criterion referencing and anonymous marking; and detailed scrutiny and sharing of assessment practice through peer review. However, none of these changes has been fully implemented in the PhD, least of all adapting assessment methods that show ILOs are tested in the final examination. Factors that may have inhibited change include the uniqueness of each candidate's research topic and experience that arguably require professional judgement of both research output and professional and personal attributes in the context of the subject. One of Denicolo's (2003:90) respondents, for example, considered that doctoral assessment was *'inevitably subjective in the end, given the complexity and originality of any thesis'*.

The cohort nature of assessment in taught programmes affords opportunities for markers and external examiners to test consistency and reliability by scanning the work of the cohort at the same time as evaluating to what extent different students have met the ILOs, a feature not available in the PhD, which also does not allow anonymous marking. While the PhD examination does not allow for cohort comparisons, the external examining system is common to all degrees in the UK including the PhD and arguably provides a form of moderation within subject groups while contributing to maintaining consistent standards. Some, however, lack confidence in the operation of the external examiner role. In the PhD this is either because of their perception of power relations and deferral of authority to the external combined with variation in universities' procedures for examination and their application, or lack of training and potential differences between examiners leading to inconsistent approaches (Morley, 2004; Denicolo, 2003; Wallace and Marsh, 2001).

The tension between different stakeholders concerning the purpose of the doctorate may also have inhibited change. For example, whereas the primary focus of examiners' judgements is on the candidate's research achievements and output (Denicolo, 2003) with assessment of professional and personal qualities evaluated implicitly, others argue that, since the PhD is no longer a degree intended solely for those aiming for research careers, its assessment should more overtly focus on attributes that enable graduates to enter wide-ranging employment destinations. The IUBMB Standards go some way towards defining both research and professional outcomes for biosciences, but as the IUBMB themselves suggested, it is difficult to argue that the degree has a common standard globally since the standards are not compulsory. Similarly, the existence of the UK-wide doctoral qualification descriptor does not guarantee that all PhD examiners have similar standards in mind when examining candidates.

Many of those who obtain PhDs in the UK move into wide-ranging careers at home and internationally. This suggests that graduates possess sufficient transferable skills to flourish in their roles and that, for some employers at least, the PhD's international status means that global standards are being met. However, UK employers (HEC, 2012) suggest that there remains a continuing mismatch between expectations concerning the intellectual capital of PhD graduates in the workplace and doctoral learning outcomes.

Chapter 3: The examination and its relationship to the modern PhD

With respect to all doctoral programmes there is an associated...history which means each programme has considerable educational inertia...Consequently, amendments or changes to the attainment process of the doctoral degree will not be easy to implement in traditional universities. Especially if retention of the historical characteristics of the doctorate is desired. (Noble, 1994:11)

1 Introduction

Concerns about fitness for purpose have resulted in suggestions that the final examination, particularly the viva, may need to be *'re-envisioned'* to reflect the purposes of the modern PhD (Nyquist and Woodford, 2000). Given the major changes to other forms of assessment in higher education explored in Chapter 2, the unchanging nature of the UK PhD examination is striking when considered alongside the transformation of the PhD and its purposes. While the *'educational inertia'* identified by Noble continues to affect doctoral assessment, a revolution has taken place in doctoral training linked to the need to prepare growing numbers of graduates for a wide range of employment destinations in a competitive international market. Its critics (e.g. Morley et al, 2002; Tinkler and Jackson, 2002; Denicolo, 2003, Carter, 2008) believe the final examination does not necessarily serve all candidates well.

In this chapter I review literature about the final examination⁶ in the context of the PhD's evolution. The relatively few empirical studies are explored in section 2. Their focus includes: the purpose and conduct of the viva; criteria used to evaluate the thesis; and the content and nature of examiners' reports. In contrast with the limited amount of empirical research, a growing body of policy publications, reviewed in section 3, explore the examination and the broader field of doctoral education, in the UK and internationally. They concern doctoral standards and outcomes, power relations, and different examination models. I also review practitioner guides for candidates, examiners and supervisors on successfully navigating the doctoral process, before turning to the fitness for purpose of the examination in the context of diverse graduate careers, and the concept of *'doctorateness'* (Denicolo and Park, 2010).

The few empirical studies concerning the viva reflect the difficulties encountered in researching this part of the examination. Three publications emanate from viva observations (Trafford, 2002; 2003 and Trafford and Leshem, 2008). Others explore examiners' judgements

⁶ By *'final examination'* I refer both to the evaluation of the thesis and the examination of the candidate in the viva.

in the thesis and conduct of the viva, using methods such as: interviews with doctoral candidates or graduates (Hartley and Jory, 2000; Wallace and Marsh, 2001; Tinkler and Jackson, 2002), examiners or supervisors (e.g. Mullins and Kiley, 2002); and questionnaires (e.g. Jackson and Tinkler, 2001; Tinkler and Jackson, 2002); or analysis of examiners' reports. Empirical work with an international focus includes several studies exploring what can be distilled from examiners' reports about their judgements (e.g. Nightingale, 1984; Johnston, 1997; Lovat, Holbrook and Bourke, 2008; Holbrook, Bourke, Fairbairn and Lovat, 2014; Bourke and Holbrook, 2013; Lovat, Holbrook, Bourke, Fairbairn, Kiley, Paltridge and Starfield, 2015). It was in the context of the empirical studies and the growing interest in the PhD examination apparent from other literature that the current study was conceptualised.

2 Empirical studies

The perception that the doctoral examination is variable, unreliable and is a recurring theme in the literature (e.g. Tinkler and Jackson, 2002). This was perhaps to be expected in the light of the changed approaches to assessment in higher education, especially since, as observed by Denicolo et al. (2000:69), several '*dilemmas*' are inherent in the examination process.

Empirical studies span several themes, including: the inconsistency of the process; the viva – its purposes, how it is conducted and the impact on the candidate; criteria used by examiners to judge the thesis, including how examiners decide what forms of knowledge it is important for doctoral candidates to demonstrate and how their individual approaches influence their decisions; and the content and nature of examiners' reports. Combining the outcomes of several research projects, including 62 questionnaire responses, Denicolo (2003) explores the diversity inherent in doctoral assessment, summarising some of the widespread anxieties regarding the process, which she describes as '*disturbing*' (p.84). While her respondents argued for examiner training and clarification of assessment criteria, importantly, Denicolo found that the standard required for '*a straight pass*' (p.89) is universally '*stringent*'. She concludes that the thesis is the primary source of evidence for the quality of the candidate's research yet recognises the importance of the viva, although of a lower priority. According to several authors, however, it is the viva that fails to meet modern standards of assessment.

2.1 The viva: its purposes, how it is conducted, and the impact on the candidate

The perceived weaknesses of the viva are accentuated by the convention for it to be held '*behind closed doors*' (Wallace and Marsh, 2001:37). Interest in the viva has gradually increased, particularly among some social science researchers, who approached it from the

perspective of a social and political phenomenon. Since 2000, policy developments have sought to regularise the conduct of the doctoral examination, yet evidence suggests these have had only a minor impact on the viva, which remains individualistic with the potential for variability, within and across disciplines. However, it remains unclear to what extent this materially affects examination outcomes.

There are two principal reasons for the limited amount of independent empirical research on the viva. First, confidentiality and the prospect of an individual being present who is unrelated to the examination in any formal capacity raise concerns for all involved. Examiners and supervisors, as well as candidates, may be nervous about how events will unfold. Above all, in all except two UK universities⁷, the viva is closed to everyone but permitted actors. Lack of access is therefore a major obstacle to independent research. The second reason is also important. Many examiners and supervisors in subjects where research into higher education is not the main focus do not question the viva process. As Poole (2015:92) observes: *'experienced academics...rarely turn their critical gaze on to doctoral assessment itself'*. This may be because this is the way it has always worked in their discipline; therefore they do not see why it requires change (Park, 2003). Alternatively, some supervisors and examiners may simply accept the viva as the second element of an interdependent, two-part process that enables judgement of the candidate's achievements. In a survey of institutional policy *'the PhD examination was uniformly defined as a two-stage assessment involving the submission of a thesis and, normally, a viva voce'* (Tinkler and Jackson, 2000:179). Taking this perspective, the process may appear more coherent, with much of the evidence derived from the thesis with the viva employed as a 'checking' device (Jackson and Tinkler, 2001; Tinkler and Jackson, 2002). Several authors, however (e.g. Tinkler and Jackson, 2004; Phillips and Pugh, 2010), identify multiple purposes for the viva over and above authenticating the thesis.

Purposes of the viva

The rationale for the viva was questioned as early as the 1960s. Data from a survey of graduate deans in the United States led Berelson (1960:199) to critique the *'defense of the dissertation'*, comparing it to the challenging oral examination that had been in place when the PhD originated in Europe. He even suggests that the viva had become *'something of a routine that both faculty and students have to get through...neither quite knows why'* (pp.199—200). One of Berelson's respondents had a more extreme view:

⁷ The Universities of Manchester and Oxford.

What was originally a day of glory, the emergence of the butterfly from its chrysalis, has deteriorated into a perfunctory hearing...this moment...should be a rite of jubilation and ...deep intellectual significance. Insofar as it is a meaningless relic of the Continental system, it might better be abandoned. (p.200)

Berelson contextualises this idiosyncratic opinion by acknowledging that most of the graduate deans disagreed with it, although almost half of them agreed the defense had become a ritual rather than fulfilling a *'useful function'*. Almost 60 years later, the viva remains part of the examination process in the UK, whether as a result of the *'inertia'* suggested by Noble (1994:11) or because it still serves important purposes, even if *'fundamentally flawed'* (Jackson and Tinkler, 2001:356).

According to Noble (1994:67), the viva is an *'anachronism'* that *'should be dropped as an examination requirement'*. Noble's data emanated from a study involving the views of 67 academics from Australia, Canada, the UK and the United States, supporting the graduate deans surveyed by Berelson, who considered the oral defence *'outmoded'*. Noble's suggestion that the oral defence at best only repeats the information provided in the thesis, also aligns with the findings of Jackson and Tinkler (2001) and Poole (2015), who both contend that the viva rarely changes the examination outcome. Only 6% of Jackson and Tinkler's (2001:360) academic respondents, however, *'regarded the viva as totally redundant'*, examples of its value being in borderline cases or instances of disagreement between examiners. One of Poole's (2015:98) respondents suggested it could have:

Surprising outcomes...[such as] a pass for a student who both examiners thought they would be failing after reading the thesis and the reverse – students failing where the examiners thought...they would be passing.

An uncontested purpose of the viva is its role in confirming the thesis is the candidate's own work (Poole, 2015). In addition to providing an opportunity to rule out plagiarism, however, the viva enables examiners to ascertain the depth and breadth of the candidate's knowledge, both in their topic and the wider field.

Berelson (1960) touches on several viva purposes, including the significance for candidates of formally becoming a scholar and the benefits of enabling faculty to exchange views about academic standards. Jackson and Tinkler (2001) explore purposes in great detail, having analysed policy documents from 20 UK universities generated for a previous publication (Tinkler and Jackson, 2000), then combined these results with questionnaire responses concerning the examination process from lecturers and candidates in two UK universities. The

authors aimed to discover the roles and significance of the viva, placing particular emphasis on its purposes. Many of the responses related to the viva's '*ceremonial, celebratory and confirmatory*' purposes (Jackson and Tinkler, 2001:360, and see Berelson, 1960). Denicolo's (2003:84) respondents considered the viva an opportunity to evaluate '*the candidate's ability to explain orally his/her work and to provide indication of her/his breadth and depth of knowledge*'. Jackson and Tinkler's (2001:361) academic respondents used the phrase '*rite of passage*' to describe the entry of the candidate to the academic community. However, none of the candidates mentioned '*rite of passage*' focusing instead on the authentication purpose and, in the case of one of those who had a negative experience, the '*misery and humiliation*' (p.362). Another candidate thought the purpose was '*to harass certain students and encourage... others*' (p.362), implying inequity. The authors concluded that candidates '*overwhelmingly*' (p.62) regarded the viva as a summative rather than a formative assessment with the main purposes as authentication of the thesis and to assess the student's knowledge and understanding.

Jackson and Tinkler (2001) suggest that the viva's purposes depend on the quality of the candidate and their thesis, with a strong distinction between '*strong*' or '*successful*' and '*borderline*' or '*weak*' cases (p.360). In the former category, 30% of academic respondents saw the viva as a formative experience that afforded candidates an opportunity '*to discuss and develop ideas with an expert in the field*', including provision of advice on publishing [from] their thesis (25%). For weaker candidates, the viva was seen by 29% as '*a forum within which examiners could provide constructive feedback and guidance*'. According to Jackson and Tinkler's respondents, the viva enables examiners to confirm their judgement of the thesis and to test their judgement in cases where they have already decided the thesis is not good enough. In the latter case respondents considered that, other than where the thesis was borderline, it was almost impossible for candidates to redeem themselves in this situation, irrespective of how well they performed in the viva.

The assertion that the viva '*can be a good forum for those with public speaking skills who are confident and who can think on their feet, all of which has nothing to do with a doctoral programme or subsequent employment*' (Noble,1994:67) does not reflect modern expectations of graduate outcomes which include a variety of professional and personal attributes sought by employers, including excellent communication. The outcome of Poole's (2015) survey of 27 academics in the field of education at 16 UK universities, including semi-

structured online interviews with seven respondents, also suggests that the viva potentially disadvantages those with poor communication skills. He argues that *'cognitive style factors... may...affect how a candidate is perceived by examiners at the viva'* (p.97) and that those who perform poorly are *'not necessarily inferior, or less doctoral in nature, and may actually be capable of offering very perceptive observations if allowed sufficient time to ponder'* (ibid).

Poole's respondents (2015:96) regarded the *'rite of passage'* concept as negative. He suggests this might be linked to its *'disagreeable connotations'* in anthropology, and that examiners who view the viva negatively might be more likely *'to approach it in a confrontational spirit'* (p.96) than those who do not. Poole's respondents displayed similar views to those of Jackson and Tinkler (2001) concerning the summative and formative purposes of the viva. They also referred to cases where the thesis was so outstanding the examiners congratulated the candidate on passing at the beginning, a practice generally disapproved of, yet 32% of Jackson and Tinkler's candidates also disclosed that the examiners had informed them of their judgement at the beginning of the viva and in 74% of cases, the viva had only confirmed examiners' original judgements.

The perception that the viva does not contribute substantively to examiners' judgements is evident in the most recent publication emanating from a 3-project research programme aimed to improve understanding of the Australian doctoral examination. Lovat et al. (2015) aimed to explore the differences between the Australian examination and those including a viva by making comparisons with New Zealand and the UK⁸. They reviewed a selection of the literature concerning the viva, observing that: it is rarely perceived as making a significant difference to examiners' decisions, referring to the practice in some cases, of informing candidates that they have passed before the viva begins (Jackson and Tinkler, 2001); and that it may serve a celebratory or ritualistic purpose. Interviews with examiners from the three countries involved in the study aimed to explore if the introduction of a viva into the Australian doctoral examination would be beneficial. Examiners appeared to favour the system they were most familiar with (Kyvik, 2014), several of the UK examiners suggesting that *'they could not imagine the process being adequate or complete without [the viva]'* (Lovat et al., 2015:15-16).

In common with Jackson and Tinkler (2001) but by contrast with Poole (2015), Lovat et al. (2015) found that the viva rarely substantively alters the examiners' judgement of the thesis.

⁸ Most of the research associated with the Australian programme concerned the content and nature of examiners' reports and is explored in section 2.4 below.

On this basis they argue that *'were the Viva an essential component in influencing the final decision...this would have manifested itself'* (p.16), nevertheless acknowledging that the viva was *'highly regarded'* (ibid) by the UK and New Zealand examiners. Examiners' reasons for wishing to retain the viva included: appreciating the opportunity it provided to meet the candidate as opposed to judging their achievements solely through the thesis; and the importance of having met them when it came to providing support in the future, e.g. by writing references. The authors also report *'minority'* (p.17) views from UK examiners, that the Australian process of appointing normally three external examiners and not including a viva was preferable to the UK system. This opinion was linked to *'examiner weariness'* caused by the amount of time needed to examine a PhD thoroughly. Some respondents thought over-reliance on the viva could arise as a result of the inadequacy or lack of scrutiny of examiner reports. Others commented on the unsatisfactory experience of candidates in a minority of cases. Nevertheless, rather than arguing for the elimination of the viva from the process, UK and New Zealand respondents thought it should act as *'an opportunity for clarification, closure or celebration that sat on top of the rigours of the report-only process of individual examiners'* (p.18). Australian examiners with experience of UK and New Zealand systems, viewed the viva as beneficial for *'closure'* (p.19) for the candidate, although agreed that it did not alter the judgement. The authors conclude by articulating some of the benefits of the viva for the candidate, not least *'involvement in the final and most crucial stage of the process with which they have been intimately involved for many years'* (p.20). They suggest both Australia and the UK could learn from one another's strengths in the examination processes. Nevertheless, they warn that the doctoral examination might be at risk from diminishing quality assurance procedures, particularly since it *'is under pressure from the combination of an increasing mass industry and reduced resourcing'* (p.20). If it is true that quality assurance of the doctoral examination is decreasing, the viva's role in providing an opportunity to authenticate the thesis, while allowing examiners to offer formative feedback that will improve the final version, should perhaps assume more prominence.

How the viva is conducted and the impact on the candidate

The conduct of the viva is not associated in the literature with its purposes but with the examination process itself and its impact on the candidate, particularly in cases where examiners are combative.

Empirical studies regarding the conduct of the viva include Tinkler and Jackson's 2002 publication using data generated for their earlier work based on institutional policy documents and questionnaire data, with the addition of new material from pre and post viva interviews with 35 candidates/graduates. The authors aimed to explore how candidates prepared for their viva given its private nature, and its '*critical importance*' in relation to '*the quality of student experience*' (ibid, p.87). Jackson and Tinkler's previous research (2001:363) suggested that candidates' experiences of the viva '*may have significant implications for how they regard their intellectual competence and...academic prospects*'. Their evidence showed that while 17% of candidates reported that after the viva they perceived their level of academic competence to be lower than before, 53% experienced '*increased perceptions of academic competence*' (p.363). Asserting intellectual competence is one of the two reasons cited by the authors for the viva's importance in the student experience, the other being that, for approximately 25% of candidates (Tinkler and Jackson, 2002:87), '*the viva is a site of decision making*'.

The authors explore how candidates' preparation was influenced by their basic skills, including the ability to communicate and perform under pressure, the content of the thesis and their expectations of the viva. They suggest that the reason few candidates access literature on viva preparation emanates from the acknowledged variability of the examination and the difficulties encountered in preparing for it. The authors conclude that '*lack of transparency about viva purposes, [and] absence of clear guidelines for its content and conduct provide the conditions for...divergent and sometimes unconstructive practices*' (p.96 and see Jackson and Tinkler, 2001). These findings support those of Denicolo et al. (2000:68) who suggest the viva's variability is linked to different examiner interpretations of their role. Some were described as '*combative*' and likened to participants in '*the Inquisition*', whereas others initiated '*collegial debate about points of interest and differences in perspective*' (p.68) and were keen to reassure candidates at the beginning of the viva on the quality of the thesis. Such variation is identified as '*inequitable*' (p.63). Denicolo (2003) also observes that some examiners perceive the viva to be confrontational, expecting the candidate to argue their case strongly and stand up to hostile examiner questions (Pearce, 2005).

The theme of rigorous and combative viva examinations is developed by other authors, including Carter (2008), who undertook a semi-structured discussion with a panel of 23 faculty from the University of Auckland. She describes doctoral examiners as '*faceless, feared and potentially testy*' actors in the process (p.365), yet her respondents suggested that examiners

were *'determined to do well at the task of examining'* (p.372), although it is not clear whether this relates to upholding standards or helping the candidate to navigate the examination successfully. However, as her discussants observed, the doctorate *'is the highest degree that the university awards... therefore the candidate should properly be put under rigorous scrutiny'* (p.371). They claimed that if a candidate became noticeably stressed, an examiner would most likely reduce pressure (Pearce, 2005), although this was not necessarily in the candidate's interest. While these nuanced approaches may be well understood by many examiners, the candidate is by definition a novice. Unless they have been particularly well prepared by their supervisors, they are unlikely to be aware of the subtleties of the situation. Murray (2009) observes that doctoral examiners are required to test the candidate's knowledge *'to its limits'* (p.40). The implication is that this situation may cause the candidate to feel under pressure, even though most examiners do not intend to be confrontational. Exploring rigour in the viva with her respondents, Carter (2008) found two polarised positions. One group suggested *'the oral was entirely benign'* (p.371), while the others were in favour of treating the viva as a formal, summative examination, since the doctorate was the highest qualification a university could award. It was even suggested that *'A vigorous battering of questions is a sign that the examiner thinks the candidate is strong enough to perform well'* (p.371). Most respondents believed that examiners should be presumed to have good intentions, wanting to support weaker candidates to achieve a successful outcome.

Access constraints limiting empirical research have perhaps increased the focus on anecdotal evidence that exists concerning candidates' viva experiences. Hartley and Jory (2000:76) identify research limitations as a reason for much of the literature on the viva being *'anecdotal'* rather than the result of direct observation. Pearce (2005:79) for example, suggests the viva *'features as the most potentially traumatic event'* in a candidate's life, an assessment echoed by several commentators, e.g. Hackley (2012), Gibney (2013) and Bassnett (2014). All mention the potential for *'rogue'* (Hackley, 2012:1) examiners to *'dismiss years of work as irrelevant, misguided or simply wrong'*, leaving the candidate traumatised by the viva experience. Hackley's assertion that *'for candidates, the viva is the biggest day of their life so far. A bad one is a shattering emotional trauma...from which some may never recover'* (p.1), sounds melodramatic but is supported by the experiences reported by some candidates, who may be *'hugely emotionally invested'* in their doctorate (Anonymous Academic, 2016:1).

Strikingly, 60% of the 88 candidates from different disciplines who contributed to Jackson and Tinkler's 2001 study described the viva as *'enjoyable'*, whereas 20% described it negatively, using terms such as *'hostile'* or *'sarcastic'* (p.362). One of Poole's (2015:95) respondents also used derogatory terms to describe the behaviour of some examiners, such as *'pompous'*, *'aggressive'* or *'arrogant'*.

These descriptions of the oral examination might suggest a theatrical event that could be construed as unregulated and eccentric. Carter and others, e.g. Murray (2009), also the PRES results, however, suggest that relatively few doctoral candidates undergo a traumatic or unfair experience. However, Pearce's (2005:22) assertion that *"problem vivas" may well be a statistical minority, but when they do occur the consequences can be very serious indeed'*, is a salutary reminder of the impact of a difficult viva on candidates.

Several authors, including Hartley and Jory, counsel against attempting to generalize from anecdotal evidence. Tinkler and Jackson (2004:127) express concern about how candidates might be affected by *'grapevine stories'*, while Trafford and Leshem (2008:201) advise:

Stories of the occasional unpleasant viva pass around more rapidly than accounts of the majority of vivas that are successful and harmonious (Major, 1994:8)...many candidates acquire knowledge about what occurs in doctoral vivas from individualised and often sensationalised second-hand accounts.

Hartley and Jory (2000) conceived their study involving 100 psychology graduates having become aware of anecdotal reports of the apparently frequent negative experiences of candidates in the viva. They developed a questionnaire based on responses from six semi-structured interviews with five academics and a recent postgraduate at one university. Their survey aimed to address the *'problem of generalisation from limited evidence'* (p.77). The authors focus on candidates' perceptions of their viva experience, recording gender, age, mode of study, the amount of preparation (including mock vivas), the extent of the candidate's anxiety and the length of the viva. The sample candidates had either been asked to undertake major revisions or had failed (one); unsurprisingly, these 16 reported negative memories of the viva. While 85% of the candidates had *'successful'* vivas, only 44% of this group found it to be a *'positive experience'* and 39% described it as *'negative'* (p.86). However, 82% of the respondents thought their viva had been fair. This juxtaposition of a significant number of the overall sample feeling negative compared with 82% thinking their viva had been fair suggests that some candidates find the experience of the oral examination more challenging than others (Noble, 1994), and/or that some examiners make the viva a more acceptable

experience than others. In commenting on the validity of their findings, which included suggestions from candidates for improving the process, Hartley and Jory offer two caveats: that the sample studied may not accurately have represented the entire population of psychology doctoral graduates; and that self-reporting has both advantages and disadvantages. The use of questionnaires enabled a larger group to be sampled, but did not allow for collection of rich data such as that generated through interviews. Neither did it enable the authors to identify confident respondents who did not prepare for the viva, those who felt traumatised by it, or to make clear they were not seeking reports of adverse experiences. They conclude it is highly likely candidates' feelings about their viva are affected by the outcomes.

Contrasting examiner approaches were identified by Wallace and Marsh (2001), who undertook interviews with six successful mature candidates in education or social sciences whose examiners had recommended outright passes or minor corrections. Their findings suggest that some examiners were courteous and confidence building, whereas others were intimidating or egotistical, one being described as '*aggressive...and rude*' (p.49). The accounts also demonstrated candidates' varying expectations and feelings about the viva and differences in their preparedness. Two out of the six reported positive, celebratory feelings, while four had negative reactions. The authors observe that the four 'problem' vivas under scrutiny presented two particular difficulties. In one case, it appeared that the candidate's emotional resilience was being tested alongside their thesis, whereas in the other three, examiner and the candidate perspectives on the field of research differed, with the examiners wanting the candidates to '*recant*' their position. Wallace and Marsh describe this as an '*inquisition*' (p.54). They conclude that it was '*the conduct of the examiners which dictated whether the viva was a positive or negative experience for the candidate*' (p.55).

When Hartley and Fox (2004) compared Wallace and Marsh's research with Hartley and Jory's (2000) study of 100 psychology graduates, they were only in partial agreement with the conclusion that a negative examiner behaviour was likely to affect successful candidates' feelings about their experience. They support Wallace and Marsh's recognition of the need to draw attention to the inappropriate behaviour of some examiners in the viva, although in revisiting their data, they found that even candidates with a positive view of both examiners might have felt their viva was a negative or neutral experience, because it was a stressful life event. They therefore conclude that Wallace and Marsh '*perhaps over-emphasised the effect*

of negative behaviour on the part of examiners' (Hartley and Fox, 2002:30). This perspective is supported by Jackson and Tinkler (2001), who found that candidates who passed, as well as one who was referred, had negative experiences in the viva.

These studies demonstrated the variability inherent in the viva, irrespective of subject or outcome, a view shared by one of Poole's (2015:94) respondents, who commented on the diversity of candidates' experiences, "*...often as a direct result of inconsistencies of approach from examiners, which seems inequitable and unfair*". Poole's interviews with 27 examiners concerning the '*fitness for purpose of the doctoral viva*' (p.93) in the UK uncovered polarised opinions regarding examiner behaviour, even among this relatively small sample in the field of education. For example, while one respondent suggested the viva was an unfair process because of inconsistent examiner approaches, another thought that the rigour arising from: scrutiny of examiner CVs prior to appointment; the presence of other actors in the viva, e.g. the supervisor; the exchange of pre-viva examiner reports; and a jointly written final report, all contributed to safeguarding outcomes. Most acknowledged the potential for inconsistency and agreed some level of subjectivity was inevitable, but overall thought the process was '*reliable*' (p.95).

The studies emphasise the complexity of the viva as a socially constructed phenomenon, including '*the power relations that inevitably apply*' (Wallace and Marsh, 2001:37), and support the view that the process lacks consistency, leaving open the possibility of inappropriate examiner behaviour (Tinkler and Jackson, 2002). However, they also demonstrate that the viva is a testing time irrespective of outcome and that to some extent, each candidate brings their individual perspective to the experience.

Only two authors have undertaken research emanating from viva observations, Trafford (2002, 2003) and Trafford and Leshem (2008). Their work stands apart because it explores the viva using their own observations, as examiner, supervisor, independent chair or candidate. The authors took notes during PhD vivas, paying particular attention to questions asked of candidates by examiners; candidate responses were not recorded. Trafford's symposium paper (2002) and related article (2003) use the data from 25 vivas⁹. The study found that the more general questions asked during the viva '*transcend disciplines*' (Trafford, 2002:2). Trafford provides examples of the questions examiners asked during two vivas, one '*successful*' (p.10),

⁹ Of which 14 were in education or education management, 4 in applied sciences, 2 in business, 1 in psychology, 1 in bio-medicine, 1 in marketing, 1 in history and 1 in computing.

the other *'unsuccessful'* (p11), from which it is possible to discern the thesis strengths and deficiencies identified by the examiners. Interestingly, Trafford (2003) concludes that, rather than displaying the negative characteristics suggested by other authors, examiners were friendly and encouraging to candidates, being particularly keen to initiate discussion with those who displayed *'doctorateness'* (p.119). Some candidates appeared disappointed that so few questions related to the thesis content. Trafford suggests that content was *'the vehicle for research and not the determinant of doctorateness'* (ibid) and therefore that examiners only explored it in detail if they were not satisfied with some element of it, preferring to challenge candidates with searching questions. He concludes that it is *'possible for candidates to anticipate patterns of questions'* (p.120). Trafford and Leshem's 2008 handbook for candidates is explored further in section 3.4.

2.2 Criteria used by examiners to judge the thesis

Literature addressing this critical question informs debates concerning standards of doctoral training and outcomes. Only three authors have attempted directly to define examiners' criteria systematically, the most comprehensive being Lovitts (2007).

Doctoral examination procedures and regulations are the main focus of Nightingale's (1984) review of 139 examiners' comments on 58 theses at Macquarie University, which covered several process-related areas of enquiry, including advice on making judgements. These included examiner anonymity and alternatives to the thesis as well as the criteria used by examiners to recommend the award of dissertations, which Nightingale views as *'inadequate'* (p.137). She summarises the sketchy guidance in PhD regulations concerning assessment criteria, which alluded simply to the standard of *'literary presentation'* (p.146), stated that the thesis should be the student's own work and provided maximum word lengths. An additional document routinely sent to PhD thesis examiners provided further guidance, clarifying that the degree was awarded on the basis of the thesis only, since it represented *'the culmination and achievement of the candidate's study and research'* (p.147). Guidance for examiners also stated that the thesis should make *'a distinct contribution to the knowledge of the subject and afford evidence of originality shown either by the discovery of new facts or by the exercise of independent critical power'* (ibid). The many ways in which candidates may demonstrate originality is summarised by others, the most comprehensive being Phillips and Pugh (2010).

In her study of 51 examiners' reports from 16 theses in the University of Canberra, Johnston (1997) distinguishes between *'a significant contribution'* and *'originality'* when reporting

institutional criteria, which at Canberra stated the thesis should: *'make a distinct and significant contribution to knowledge or understanding; and afford evidence of originality shown either by the discovery of new facts or by the exercise of independent critical thinking'* (ibid, p.335). Johnston's distinction between the two concepts reflects intra- and inter-disciplinary differences. Poole's (2015) respondents in the field of education demonstrated that it was difficult to develop a common definition of originality, even within the same field. Poole also used the term *'original contribution'* as well as *'originality'*. Three of his contributors held divergent views about the interpretation of originality. One suggested that examiners developed the ability to make a judgement about originality through experience. Another, when asked if judgements on the question of originality could be *'idiosyncratic'* (ibid, p.100), agreed this could potentially be the case, especially if examiners did not pay attention to guidance or if guidance was poor. The third respondent surprisingly suggested they had *'never had a conversation with anyone either in a viva or outside about originality'* (ibid). Poole uses these views to argue that examiners make assumptions about what constitutes originality and that this calls into question the reliability of the viva.

This fundamental requirement for achieving a doctorate has changed little or not at all since Nightingale's study. Having compared the Macquarie criteria with those at other Australian universities, which mentioned either a *'significant'* (Monash) or *'substantial'* (Queensland) contribution to knowledge/ learning in the field or disciplinary area (p.148), Nightingale concludes that such criteria were *'an insufficient basis on which to evaluate the work of several years and on which to distinguish between levels of achievement that may possibly determine the course of a whole career'* (p.148). Nightingale obtained access to discipline-specific examples of criteria for judging theses, including education and pharmacology. These demonstrated an emphasis on clarity of expression and theoretical position in the field of education, whereas in pharmacology emphasis was placed on comprehension of the field, appropriateness of experiments to investigate the problem, and the ability to communicate scientific findings. This suggested different priorities among subjects regarding the most important components of a thesis. In conclusion, Nightingale argues for clarification of research degree awarding criteria, including guidance accessed by examiners, yet acknowledges the difficulty of defining more explicit statements of achievement. Denicolo (2003:86) also observes the generality of the criteria commonly used for achieving a doctorate, citing *'an original contribution to knowledge and having the potential to be published in some form'*. Denicolo found that the criteria of originality, contribution to knowledge and scholarship were universally agreed yet, as would be observed later by Lovitts, *'participants*

had experience of widely different, but usually implicit, ways of judging or operationalizing these concepts' (p.88).

Research that aimed to define how examiners make judgements about the quality of the thesis was conducted by Mullins and Kiley, (2002). They interviewed 30 examiners across science, mathematics/ engineering, social science and humanities, who had all assessed *'the equivalent of at least five research theses over the last five years'* (p.369) in Australia. They took account of: the stages examiners go through; the criteria they use; what influences them; the evidence they seek; and the critical points as they come to a judgement. Regarding the influence of the examiners' *'methodological paradigm'* (p.375), Mullins and Kiley found that in the sciences, examiners have coinciding views about what is important, namely: *'good science'*, summarised by the authors as *'a pertinent literature review, clear hypothesis, do-able problem, sound data analysis and methodology, and justifiable conclusions'* (p.375). The authors judge humanities and social sciences examiners, on the other hand, to have more flexible objectives, agreeing that candidates should be consistent and demonstrate *'that they had actually done what they said they were going to do rather than adhere to a particular paradigm or methodology'* (p.375). The differential approaches between science and social sciences/humanities examiners resonate with Nightingale's (1984) findings.

Mullins and Kiley's examiners differentiated between theses of different quality (Lovitts, 2007). Some, for example, described characteristics of a *'poor'* thesis, alluding to *'sloppiness'* (Mullins and Kiley, 2002:378), including typographical errors or other careless mistakes, which they thought indicated lack of rigour generally on the part of the candidate. Respondents also cited *'lack of coherence; lack of understanding of the theory; lack of confidence; researching the wrong problem; mixed or confused theoretical and methodological perspectives; work that is not original; [and] not being able to explain...what had...been argued'* (p.378) as characteristics that might be found in a poor thesis. When describing passable theses, on the other hand, examiners were likely to refer to *'scholarship...originality, coherence...autonomy and independence'* (p.379). Of primary importance was that the candidate had presented a logical argument in a well-structured, *'substantial'* (ibid) thesis and demonstrated the ability to reflect on their own work and its place in the field. As remarked upon by Lovitts (2007), examiners used *'artistic metaphors'* such as *'élan...elegance...creativity'* (Mullins and Kiley, 2002:379) to describe theses they found outstanding. An examiner in humanities described the thesis range as follows: *"you can't fail it but it's not dazzling"* (passable); *"makes you see*

an area...you thought you knew in a way that you hadn't thought about" (outstanding) (p.380).

Mullins and Kiley query the need for a doctoral viva, given that most examiners expected the thesis to pass. Yet some of their respondents viewed it as a formative assessment where they could give candidates advice, for example, on improving the finished thesis (Johnston, 1997), or on additional literature. Denicolo (2003:87) observes that there is disagreement among practitioners about *'whether the viva is a formative or summative...assessment or a mixture of both'*. Mullins and Kiley's respondents suggested candidates saw the doctoral examination as a summative event. Other authors (e.g. Jackson and Tinkler, 2001) support this finding and observe that for some candidates, the viva represents a distinct change of status from student to independent researcher and equality with academic peers. The doctoral examination clearly serves multiple purposes (Phillips and Pugh, 2010) and although formative assessment is important, the examination has a summative outcome. Mullins and Kiley conclude that assessing a thesis is complicated (Nightingale, 1984), that examiners are seeking numerous characteristics (Lovitts, 2007), and therefore that the assessment could not be reduced to a numerical marking system: *'it seems that it is not possible to "mark" each one out of 10, total the results and declare a thesis passed or failed'* (Mullins and Kiley, 2002:383).

Another study of particular relevance was undertaken by Lovitts' (2007), who aimed to define the *'implicit'* (p.xi) criteria used by in the US doctoral examiners to judge the dissertation. Based on a survey of 276 faculty in 74 departments across nine universities, it included 10 subjects in sciences, engineering, social sciences and humanities. Lovitts aimed to identify examiners' expectations and in particular to make them *'more transparent to graduate students...thereby helping them achieve higher levels'* (ibid). Her study was set in the environment of reform of doctoral education in the United States over more than a decade, when standards were a major concern. Lovitts suggests that definitions such as the requirement for the candidate to produce *'original'* work or make a *'significant contribution'* are *'largely unexplicated and mysterious'* (p.3). She observes that, at the time: *'virtually no research'* existed in the United States *'on the standards used by faculty to judge [doctoral] dissertations'* (p.4), also suggesting that *'differences among faculty'* (p.117) are pertinent, that is, if *'high-PhD-productive faculty and low PhD-productive faculty'* employ different standards and in particular whether they adjust their standards of examination from candidate to candidate.

Focus groups consisting of faculty from each institution were asked to rate the importance of six characteristics of a dissertation: *'independent contribution, originality, significance, substantial time commitment, length of document and publishable or source of publishable material'* (ibid, p.5). Their responses formed the primary data and therefore these characteristics shaped Lovitts' definitions of dissertation quality.

The most important characteristic identified was *'independent contribution'*, with *'originality'* and *'publishability'* jointly rated second. Lovitts distinguished between originality and a significant contribution, noting that these terms could be discipline specific (Delamont et al., 2002) but that no study had *'analysed these concepts along disciplinary lines'* (Lovitts, 2007:10). In exploring the *'universal qualities'* (p.27) of dissertations, Lovitts reports that faculty in all disciplines used the term *'original contribution'* (p.31) to identify new work demonstrated anywhere in a dissertation. This conflation of originality with a contribution to knowledge by her respondents on one hand suggests the close correlation between Lovitts' original characteristics, in particular *'independent contribution...originality...[and] significance'* (ibid, p.5), yet as we see in the current study, does not help to differentiate between the varied use of these terms (Clarke and Lunt, 2014).

To contextualise two emerging themes, Lovitts (2007) uses the Australian and UK literature on the topic of how examiners judge the thesis. The first is that examiners' criteria fall into *'technical'* and *'indeterminate'* qualities (p.7). The second, as already mentioned, is the expectation of originality, or a contribution to knowledge. Lovitts defines technical qualities as attributes that can be formally taught and objectively assessed, together with qualities that reflect attention to detail and affect the presentation of the dissertation. Indeterminate qualities on the other hand she defines as *'tacit'* (ibid), examples being *'intellectual grasp, coherence and critical thinking'* (ibid). Lovitts argues that strength in one area can compensate for weaknesses in others, although her respondents emphasised that weaknesses should be minimal. In describing how the dissertation demonstrates a range of achievements that cannot easily be quantified, Lovitts reflects on the oral defence. Even though by focusing on dissertation judgements, Lovitts' research implies the pre-eminence of the thesis in the examination, she suggests that in an assessment of such complexity, dialogue between the assessors and the candidate is essential so the candidate can *'clarify questions asked and answers given'* (p.23). This unambiguous support for a two-part doctoral examination contrasts with earlier research, especially Berelson (1960) and Noble (1994).

Lovitts, (p.9), observes some consistency between examiners in defining the characteristics of 'good quality or passing' and 'poor quality or failing' dissertations. Clarke and Lunt (2014) also explore the consistency of examiner criteria in a paper that includes emerging data from the current study. We devised a typology of categories emerging from responses of examiners, independent chairs and supervisors concerning the characteristics sought in the thesis and viva. While based on a small sample (n=17), the responses suggested similar 'core' principles that co-exist with disciplinary priorities, indicating subtle differences between laboratory-based and text-based subjects. Examiners appear to use the concept of originality in some form, together with engagement with the literature and other scholarly characteristics such as innovative research design and coherent arguments, to differentiate between borderline theses and those of acceptable standard.

Lovitts' respondents viewed the dissertation as a process as well as a product and identified three purposes: providing training for the candidate; facilitating their learning and knowledge acquisition; and demonstrating that they were now capable of working as an independent researcher, the first two being formative, the latter summative assessment. They also saw the dissertation as a product that provided the candidate with entry to their profession. Lovitts summarises the characteristics inherent in dissertations of four levels of quality: outstanding, very good, acceptable and unacceptable. She notes that the faculty 'balked' (p.41) when asked to describe dissertations in the unacceptable category and that her respondents 'rarely, if ever, failed a dissertation' (ibid). Faculty in the focus groups were at 'highly ranked universities' (p.39), which suggested their standards and expectations were high. The majority of dissertations the faculty had experienced were in the 'very good' category, although these candidates were unlikely to demonstrate excellence across all criteria. Lovitts' respondents also used 'artistic metaphors' (Mullins and Kiley, 2002) to describe the outstanding thesis, for example, 'creative', 'elegant', etc. Lovitts uses her respondents' criteria to differentiate between the four quality levels for each of the six components to create 'rubrics' (Lovitts, 2007:49; 53—58;97—98), or doctoral learning outcomes, which enable faculty to discern differences in the quality of dissertations and the characteristics of their authors. She intends that the rubrics should make explicit to candidates the standards expected, thereby improving their dissertation performance, although it is unknown whether the rubrics have influenced practice in doctoral training.

Lovitts' study exemplifies the complexities of evaluating the achievements of doctoral candidates and the many influences on examiners' judgements. It aimed to increase

candidates' understanding of the examination process and encourage examiners to articulate the criteria on which their judgements were based, addressing two of the major criticisms of the PhD examination, renowned for its secrecy and impenetrability.

2.3 The content and nature of examiners' reports

Empirical studies in Australia emanating from the analysis of examiners' reports present some interesting arguments for and against a thesis and viva system such as those in the UK, New Zealand and the United States; and a public or private defence.

Aiming to gain insight into examiners' expectations of candidates, Johnston (1997) analysed 51 examiners' reports relating to 16 doctoral theses at the University of Canberra. Like Mullins and Kiley (2002), she found that examiners were irritated by typographical, grammatical and spelling errors in a thesis and that poor presentation was distracting. Both studies demonstrated that examiners seek coherence and focus in the thesis. Another of Johnston's findings was the propensity of examiners to offer formative feedback (Lovitts, 2007; Holbrook et al., 2014). Johnston suggests the detail afforded by a written report enables the candidate to gain a deep understanding of the process. One of Johnston's examiners suggested that the thesis they had examined should be treated as a '*sound draft*' as it had not met the threshold level for doctoral work, but that '*with further development*' (Johnston, 1997:338), it could, and as such was recommended for a pass subject to amendments. This example illustrates how examiners use formative feedback to ensure the final thesis meets doctoral standards. However, Johnston is concerned that this examiner's comments are indistinguishable '*from many associated with the recommendation of resubmission and re-examination*' (ibid), arguing that this demonstrates inconsistency in the treatment of similar quality theses. Johnston concludes '*that most examiners have relatively clear ideas about the standards required of doctoral theses in their field*' (p.343). Some contest this hypothesis (e.g. Morley et al., 2002), and believe such an implicit approach is unacceptable. The corollary of Johnston's argument is that the academic community in any field can uphold standards through their shared understanding (and see Lovat, Holbrook and Bourke, 2008; Holbrook, Bourke Fairbairn and Lovat, 2014).

A three-project Australian research programme¹⁰ that began in 2003 has generated around 12 publications. Data are based on selective sampling of 2,121 of examiner reports (one of the

¹⁰ Based at the Centre for the Study of Research Training and Impact at the University of Newcastle, New South Wales.

later papers suggests the sample increased to 2,700), relating to 804 PhD candidates from eight Australian universities, chosen for varying research intensity. Using mixed methods, the aim was to gain an understanding about doctoral standards and the criteria used to judge thesis quality. Australian examinations were the focus of the first two projects whereas in the third, comparisons were made with New Zealand and the UK (Lovat et al., 2015), with the objective of comparing doctoral examination models to determine the purposes of the viva, which is uncommon in Australia. The remaining publications explored in this section contributed to the first two projects in the programme and are the most relevant to this study.

In the first of two earlier studies, Holbrook, Bourke, Lovat and Dally (2004a) aimed to identify categories of comment in the 2,121 reports and how these related to the examiners' assessment criteria and standards. They employed four categories established by Holbrook and Bourke (2004): examiner and process (how examiners approached the task); assessable areas covered (elements of the thesis judged by examiners); dialogic elements (how examiners engaged with candidates through reports); and evaluative elements (judgements of validity, worth and contribution of the thesis). The authors examined high and low quality theses and found that in higher quality theses, 20% of the report was devoted to positive summative comment, compared with only 2% in those of low quality, where contribution to knowledge or significance of the research was rarely mentioned. Importantly, the authors found that if a thesis was considered marginal, lack of originality appeared *'to signal the dividing line between an acceptable and an unacceptable doctoral standard'* (Holbrook et al., 2004a:113). They conclude that examiners wish to demonstrate how a thesis falls short of their expectations and to improve quality, finding it is possible from the comments to distinguish clearly between high and low quality theses, but more difficult to identify those in the middle ground. In the second study (Holbrook et al., 2004b), they analysed the core text of examiners' reports on 101 theses from one university, identifying six areas commented on by most examiners: scope, significance and contribution of the thesis, including publications; approach or method; engagement with and treatment of the literature; analysis undertaken; and the candidate's communicative competence (p.142). An interesting result of this study was that examiners appeared most engaged with *'the analysis and interpretation of the findings'* (p.143).

In another study to explore how examiners decide what forms of knowledge doctoral candidates should demonstrate, Lovat et al. (2008), assumed that examiners in a particular field *'share a familiar set of common-sense understandings about research at the PhD level and what is acceptable'* (p.66) and that these could be determined from their reports

(Holbrook, et al., 2004a, 2004b). Reports were analysed using Habermas's (1972, 1974) theory, which proposed that knowledge in any field is developed, shared and negotiated using one of three *'ways of knowing'* (Lovat et al., 2008:67): *'empirical/analytic'* concerning *'content [and] facts and figures'*; *'historical/hermeneutic'*, relating to understanding, interpreting or communicating knowledge; and *'critical or self-reflective'*, indicating freedom to draw new conclusions about *'the content and meaning of knowledge'* (ibid). In applying Habermas's theory to their sample of reports, the authors paid *'special attention'* (p.72) to the top 5% of the 804 theses, focusing in particular on examples of formative and summative comments. Formative comment was found in only 63% of reports, whereas unsurprisingly 97% contained summative comments. *'Explicit negative comment'* (p.73) was present in 46%. Given that less than 1% of the reports recommended failing the candidate, these outcomes may be partially explained by the formative comments of examiners wishing to improve thesis quality.

In their analysis, Lovat et al. use the four categories of examiner comment identified by Holbrook and Bourke (2004) in coding the data, also taking account of report organisation (structure). Theses of a higher quality are praised for: the significance and challenges of the research undertaken; the amount of work done by the candidate; appropriate use of the literature in framing the study; and the potential application of the findings. These contrast with those on lower quality theses, which include *'a preponderance of instructive comment, a relatively greater proportion of negative judgement and an... emphasis on editorial errors and inaccuracies'* (Lovat et al., 2008:71).

Using Habermas's theory, Lovat et al. (2008) conclude that even with high achieving candidates, examiners mostly see themselves as *'experts'* (empirical/analytic) with the candidate being *'subservient'* (p.72) and this leads to a technical and negative bias. Even in some cases where candidates were awarded an unconditional pass, the authors identified critical terminology, including *'shortcomings'* and *'weaknesses'* (p.73). Other comments, however, demonstrated a *'partnership'* approach (historical/hermeneutic), the relationship perceived as *'more even-handed'* (p.72). Very rare were examples of *'intellectual equality'* (ibid) (critical/self-reflective), or role reversal: *'you wish that you had written [the thesis]'* (p.73 and see Johnston, 1997). According to Lovat et al., their findings raise the possibility that examiners with a negative bias may allow it to influence their behaviour in the viva and by doing so diminish the candidate's experience, whereas those who take a *'partnership'* or even *'intellectual equality'* approach might make the viva experience more rewarding for the candidate. The quality of the candidate's thesis, being the principal focus of examiners'

judgements, inevitably influences their approaches to both elements of the examination and the content of their reports.

With the aim of testing the degree of consistency among examiner judgements, Holbrook, Bourke, Lovat and Fairbairn (2008) analysed the data set of 2,121 reports. They also used the four categories of examiner comment identified by Holbrook and Bourke (2004) to test consistency of judgements, finding that less than 2% (37 out of 2,121) of examiners' judgements showed discrepancy with others who examined the same thesis (Holbrook et al., 2008:45). The authors conclude their findings demonstrate standards are applied consistently, that the consistency and stability shown by the results are '*remarkable*' (ibid) and that standards can be identified from examination reports. They suggest it would be possible to develop indicators that would enable stakeholders, including examiners, to differentiate between '*theses of threshold quality and those of higher and highest quality*' (p.46).

Another project explored whether the qualities examiners sought were different, depending on research degree level and, more relevant to the current study, if some criteria were more important than others in judging thesis quality. Bourke and Holbrook (2013) surveyed 353 PhD and 74 research Masters examiners from five Australian universities of different research intensity. They employed 12 indicators that had been established in coding the main set of 2121 examiner reports, divided among: contribution (3); literature review (3); approach and method (2); analysis and results (2); and presentation (2). Having undertaken statistical analysis on the quantitative data, the authors contend that '*originality*' or '*contribution*' was equally sought in Masters and PhD theses, but examiners rate more highly the contributions to knowledge of PhD candidates. They conclude that before progress can be made in developing guidance '*to increase the validity and reliability of thesis assessment*' (p.415), further work is required on how a viva might affect assessment outcomes.

A later study aimed to establish '*the focus and substance of* the formative comments to candidates exemplified in examiners' reports, based on 345 examiners' reports in science and 140 in education from the main data set (Holbrook et al. 2014:1). Both summative and formative comments were identified. Summative remarks pertained to all or much of the thesis, whereas the formative, or developmental comments were judged to fall into three elements: '*identification of the concern or problem, engagement with the material pertinent to it, and direction or ideas for improvement*' (p.8). The authors found that some comments were prescriptive, for example, a suggestion or instruction to the candidate about how to put something right, whereas others consisted of a commentary intended to improve the

candidate's understanding. They undertook quantitative analysis of the formative comments, which they categorised under three headings: *argument* (including reasoning, coherence, completeness, and depth and sophistication of thinking); *project* (including theory, literature review, data quality and analysis); and *fundamentals* (including methods and thesis presentation). They also calculated the amount of formative comment in all the theses, estimating that it constituted around 30% of the content in both science and education reports.

Further analysis of examiner comments in each category established correlations between formative comments and examiner recommendations, with science examiners providing more formative comments on analysis and reporting of research, and education examiners placing greater emphasis on coverage of literature. Comparing formative comment in education and science reports, Holbrook et al. (2014) suggest that disciplinary differences are important in relation to the candidate's development as a researcher. A striking finding is the identification of the close alignment of formative and summative comments, especially when formative feedback is used to bring the thesis to the level required to pass the examination. The authors conclude that doctoral training should focus on the attributes sought by doctoral examiners and that universities should use recommendations to maintain standards. They comment on the balance of '*the pedagogical with the pragmatic*' (p.16) when examiners make recommendations for revising the thesis to a 'passable' level and suggest the candidate's learning continues when revising the thesis according to examiners' comments.

Exploration of the empirical literature emphasises the potential for variability in the examination, the challenges inherent in researching the viva and some polarised opinions. For example, while some authors seem convinced of the need for change to remove variability, others demonstrate some consistency of examiner standards in judging the thesis and the candidate, together with subtle disciplinary differences.

3 The examination's fitness for purpose in the modern context

'*There has always been something dynamic and evolving about the doctorate and its nature*' (Wellington, 2013: 1490). This dynamism is reflected in the growing body of eclectic literature on the doctorate and its purposes, particularly concerning the final examination, which explores its wider context, including the purposes of the PhD and the nature of 'doctorateness'. Many of these authors have considerable experience of the examination,

either as an examiner, supervisor, candidate, or independent chair, so their opinions are grounded in experience as well as in some of the empirical studies.

The literature reviewed below is divided into five sections. First I explore early proposals for changes to the PhD and its examination (ESRC, 1987; CVCP, 1988). Next is an analysis of publications that treat the examination as a social construct (e.g. Baldacchino, 1995; Morley et al., 2002), while 3.3 focuses on international comparisons of the examination (e.g. Hartley, 2000, Pearce, 2005; Kiley, 2009). Guides aiming to support candidates and practitioners in negotiating the PhD and its examination, some highly critical, are explored in 3.4. Literature questioning the fitness for purpose of the modern PhD and its examination is explored in section 3.5, where I address the perceived narrowness of doctoral training and its outcomes (e.g. Nyquist and Woodford, 2000; Nyquist, 2002; Borkowski, 2006; HEC, 2012) and the need for doctoral graduates to possess attributes that equip them for professional roles. Related themes are: what is being assessed in the examination, the thesis or elements of the candidate's training process and abilities (Morley et al., 2002; Denicolo, 2003; Lovitts, 2007); and attempts to define 'doctorateness' (e.g. Denicolo and Park, 2010; Wellington, 2013).

3.1 Doctoral standards and outcomes

Literature concerning doctoral standards and outcomes identifies concerns suggesting the examination process may no longer meet current purposes of doctoral assessment or deliver the necessary outcomes. Attempts were made as early as the 1980s to introduce incremental changes as a result of concerns regarding the quality of training and thesis standards. Even in a relatively small higher education sector there was awareness of international competition for PhD applicants and the need to sustain standards of doctoral education. Informed by intelligence from the UK Research Councils, the Committee of Vice-Chancellors and Principals' (CVCP)¹¹ report on 'The British PhD' (1988) recognised changes in PhD education and acknowledged some variation in training standards.

Having consulted 68 UK universities using a questionnaire to determine views on a taught component in the PhD, the CVCP (1988:1) summarise the '*inter-related purposes*' of the PhD from a recent report (ESRC, 1987), as making a contribution to knowledge and providing research training to prepare graduates for independent scholarship. They note the ESRC's insistence that '*a grounding in research techniques should be an integral part of the PhD*'

¹¹ The current equivalent of this organisation is Universities UK: <http://www.universitiesuk.ac.uk/>

(CVCP, 1988:2). With the condition that a significant research component must be retained, the CVCP recommend that all doctoral programmes should contain *'taught elements'* (p.3), to: improve the quality of the thesis; avoid isolation; and deepen and broaden knowledge of the field. These were admirable aims that brought together research and professional and personal considerations, many years before the Roberts initiative (Roberts, 2002) and the UK Research Councils Joint Skills Statement (QAA, 2004) shifted the research training emphasis to professional skills. The second topic the CVCP explored was the question of how to address the

Enormous range of achievement represented in the award of the doctorate and the concern of both the academic community and employers that the diversity of quality in doctoral theses was not reflected in the result (CVCP, 1988:4).

Given that when the CVCP report was published, the major expansion of higher education had yet to occur and therefore the PhD graduate population was relatively small, it is surprising that their members already perceived such a range of achievement. The CVCP's recommendation to introduce a PhD *'with distinction'* went as far as to suggest that *'about ten percent of PhDs awarded might fall into this category'* (p.4). It has not been possible to ascertain why no attempt was then made to introduce a graded PhD in UK universities, but it was clearly a controversial proposal. More recently, Denicolo and Park (2010) suggest there is still *'an interesting debate to be had over whether or not the UK doctorate should be graded, rather than simply having a pass or fail result at the final stage'* (p.3).

Following the CVCP's intervention, there was something of a hiatus until the 1990s, when several academic and policy studies emerged in the UK and also in the United States (e.g. Becher, Henkel and Kogan, 1994; Noble, 1994), culminating in the Harris report (HEFCE, 1996). These publications were principally driven by the need to improve doctoral training and output standards. Their influence contributed to the inclusion of doctoral degrees in the quality assurance developments explored in Chapter 2.

3.2 Power relations: the examination as a social construct

The political tensions inherent in academic life and the interests of individual participants were concerns of Morley et al (2002) and Pearce (2005), who characterise the examination, particularly the viva, as an idiosyncratic, socially constructed event. According to Pearce *'candidates, supervisors and examiners are all potentially guilty of playing a political game'* (p.27), and *'abuses of power'* (p.29) will continue unless more than two examiners are involved in the viva.

Similarly, Baldacchino (1995:71), suggests that *'the doctoral exercise is itself a manifestation of a social practice'* that should be subject to scrutiny. Having argued that the candidate is the expert in their topic, he contends that some examiners use their position of authority to confront them on their knowledge or data interpretation, potentially altering the outcome of the examination. He also suggests that examiners use the viva as an opportunity to 'show off' to the other examiner(s) as a way of embellishing their reputation. Denicolo et al. (2000) refer to *'the influence of power'* (p.65) that may encourage examiners to pursue their own ideas rather than to accept the candidate's perspective on the topic, while Baldacchino (1995), Morley et al. (2002), and Poole (2015) all explore the social interaction that takes place. Describing the viva as *'inquisitorial'* and referring to *'the aggression, sarcasm and hostility'* experienced by some of Hartley and Jory's (2000) respondents, yet simultaneously acknowledging the advantages afforded by the viva, including the opportunity to ask questions on the thesis and to understand if the candidate led the research, Morley et al (2002:266) contend that *'in the power-laden micropolitics of the academy, many diverse interests are...invested in the viva. It is almost an academic equivalent of debutantes being presented at court'*. They conclude that a few actions would address the weakest elements of the examination process, including: more guidance for candidates; some agreement about the balance between the thesis and viva in making the judgement; more transparency and accountability in the selection of examiners; and the recording of viva proceedings by universities. Since these suggestions were made, advice available for candidates has increased and some universities either use independent chairs or record vivas. Selection of examiners and the balance between the thesis and the viva in the final judgement, however, have not been addressed.

3.3 Comparisons of doctoral examination models

Comparisons between the UK and other systems are normally made as part of an argument that identifies one or other process as superior in those the author has experienced. This subsection summarises publications that as their main focus aim to compare one system with others.

The closed nature of the viva is the issue that generates most concern in the UK, although as Hartley puts it: *'If you thought having a viva in the UK was complicated enough, you might be surprised by how it is done in other countries'* (Hartley, 2000:22). Hartley describes the viva process in 19 countries but does not express a preference for any one model. Morley et al.

(2002) suggest that some of the features of the US system are commendable, in particular dispensing with the confidential nature of the oral examination, whereas Lovat et al. (2015) conclude that elements of all three systems they compare (Australia, New Zealand and the UK) have merit. Referencing the US system, both Morley et al. (2002) and Pearce (2005) call for the UK's viva model to be replaced by a panel of examiners, within which there is no *'all-powerful voice'* (Morley et al., 2002:269) – that of the external examiner. Kiley (2009) also raises the question of the weighting of panel members' opinions in the assessment process, including whether the external examiner's view should take precedence. Pearce (2005) advises prospective examiners that until the viva begins it is not be clear which of them will take the lead role.

Other authors compare the format of the examination in the UK and other countries. Denicolo (2003:87), touches on *'the very different practice in European and North American countries in which public vivas are the norm'* but does not suggest that her respondents favoured these alternatives. Pearce (2005) compares the UK examination process unfavourably with other approaches, suggesting that in France *'the PhD has maintained its reputation for truly cutting-edge intellectual engagement of the highest order'* (p.13). Her assessment of the French model is that the experts who comprise the PhD examination panel in France mean it is more like a *'trial'*. She suggests that the public defence model *'becomes a very different affair in such a public arena and...points to some of the problems with the watered-down UK version'* (ibid). Pearce also explores doctoral examinations in Finland, Sweden, Australia, the United States and Canada, claiming that these *'alternative'* examination processes serve to emphasize *'just how closed and inscrutable the UK model is'* (p.16).

Using data from an evaluation of PhD training in Norway in 2010 involving international members of PhD assessment committees, Kyvik (2014) mainly focuses on inter-country comparisons of assessment procedures and includes exploration of whether the viva should be public or private, which many of his UK and US respondents commented on. He found mixed opinions: *'many were positive to the public character of the defence, but just as many were critical'*. (p.147). Those in favour of public defences cited: fairness to the candidate; the opportunity to test the candidate's ability to defend their research; and the *'rite of passage'* role (Tinkler and Jackson, 2004; Lovat et al., 2015). UK respondents, however, were opposed to public vivas. They argued the examiners' decision is made on the basis of the thesis and presumably this would obviate the need for a public demonstration of capability, also that the viva as a ceremonial occasion would prevent them from asking searching questions of the

candidate arising from reading the thesis. They alluded to the potential for public humiliation of the candidate if the defence did not go smoothly. Kyvik's respondents generally eschewed the theatricality of the public event in favour of the more informal approach in the UK viva that one respondent suggested *'serves the candidate and the process well. The questions can be much more detailed and the answers...more thoughtful'* (Kyvik, 2014:148). They agreed the viva afforded the ability to ask searching questions, suggesting that in a public setting examiners had limited opportunities to *'push'* or *'interrogate'* the candidate (ibid). Kyvik concludes that examiners tend *'to prefer their own national assessment procedures'* (p.149) and questions whether it would be timely to attempt to standardise PhD assessment practice internationally. Conversely, one of Poole's respondents suggested this would be *'very challenging'* (Poole, 2015:95).

In her paper on the Australian examination, Kiley (2009) explores the purposes of the doctorate and what is being assessed. Having made arguments for and against an oral examination in Australia where doctoral assessment had become a *'vexed'* issue (p.36), Kiley does not favour the supervisor acting as internal examiner because of the need to avoid *'ethical dilemmas'* and conflicts of interest (p.37). She contends there is little evidence that the examination process is improved by the involvement of external examiners (all Australian PhD examiners are *'external'*) and instead suggests that in every doctoral assessment one examiner should *'be from outside the academy'* (p.37), a potentially controversial idea given the elevated status of the PhD in the array of academic qualifications. Kiley concludes that, given the *'valuable learning outcome for candidates, supervisors, examiners and institutions'* (p.40), and in particular the opportunity for formative feedback to candidates afforded by meeting the examiners face to face, the introduction of an oral examination in Australia could improve the process.

3.4 Guides for candidates and practitioners

The increasing number of handbooks and guides concerning the UK doctoral examination reflects its perceived opacity and the need for demystification (Tinkler and Jackson, 2004:4), as well as the growing numbers of candidates. Those explored below adopt a practical rather than a policy approach. Authors aim to support candidates and other actors before and during the doctoral experience.

How to get a PhD – a handbook for students and their supervisors (Phillips and Pugh, 2010) advises candidates how to approach a doctorate, focusing on the most important philosophical

and practical steps towards completion, including: how to conduct research; the writing process and how to construct a thesis; how to avoid becoming isolated; and ways of interacting with supervisors. The definitions of 'originality' and how it may be demonstrated are comprehensive. As the authors observe, '*The main problem is that there is little or no discussion between students and their supervisors of what constitutes originality in the PhD*' (p.70). They suggest supervisors are so experienced in this area that they realise the candidate needs only to show '*an incremental step in understanding*' (ibid) to achieve originality, yet they rarely convey this to students. They contend that candidates' gradual realisation as they progress through the PhD that they have indeed produced original work is an indication of their '*academic development*' (ibid), and results from their own transformation as a researcher.

Stepping stones to achieving your doctorate (Trafford and Leshem, 2008), another handbook for candidates that focuses broadly on the doctoral context, advises that to be successful it is essential to prepare for the viva from the beginning. Expanding on Trafford's publications (2002; 2003), the book emerged from a series of workshops for doctoral candidates from 50+ countries in around 30 disciplines, combined with the observations of approximately 100 doctoral vivas in different universities. The authors obtained access '*by being either a candidate, examiner, supervisor...or Chair of the viva*' (Trafford and Leshem, 2008:7). Apart from the current independent study, Trafford and Leshem's work is the only example of an empirical study on the viva that includes observation as a method.

The book addresses engagement with literature, research design, and structure of the thesis. Two chapters are dedicated to the viva. The first encourages candidates to consider how examiners conceptualise the viva by providing examples of how they focus on the strengths and weaknesses of the thesis. It explores the distinctiveness of the doctorate as a qualification and introduces the concept of 'doctorateness'. The authors developed a quadrant model based on their interpretation of the '*relative significance of questions*' in a viva, where A represents the '*Technology of the thesis*', B '*Theoretical perspectives*', C '*Practice of research*' and D '*Demonstrating doctorateness*' (p.20, Figure 2.2). The purpose of the quadrant model is to illustrate how the quality of the candidate's thesis can demonstrate the attribute of '*doctorateness*', leading either to a '*good*' or a '*poor*' viva (pp.189–190).

How to survive your viva (Murray, 2009) is another handbook aimed predominantly at supporting candidates, in this case to navigate the viva successfully, rather than addressing the

whole approach to the PhD. It provides a concise summary of viva purposes: *'To find out: (1) exactly what went on in the research; (2) exactly what the thesis means; and (3) exactly what the student knows'* (p.xiii). Murray advises the candidate to ensure they are familiar with the rules for the doctoral examination in their own institution, emphasising the differences in practice that exist while arguing that in every department and discipline, there should be some consistent practice. The book provides insight into the many forms the viva can take and how candidates may feel during and after the event. The author addresses the viva's role in the doctoral process, is pragmatic about the potential for variation and variability and suggests that *'the concept of standard practice [in the viva] is not universally established'* (p.23). This is an unwelcome message for doctoral candidates, but an important one. Chapter 5 *'Questions'* steers the candidate through the mechanics of the viva, containing advice about preparation for the event, including advocating mock vivas (Hartley and Fox, 2004). Chapter 6, *'Answers'*, advises candidates to: focus on the question asked; use subtlety; be specific; to showcase elements of the research; and not to assume they are stating the obvious. Murray also advises the candidate to avoid emotional responses to examiners' questions or taking a *'[them-and-us] attitude towards the doctoral examination'* which *'may be fuelled by hearsay and anecdote'* (p.40), counselling against being defensive or thinking *'the examiners are out to get you'* (p.149).

The Doctoral Examination Process (Tinkler and Jackson, 2004), a handbook for examiners, supervisors and students, draws on data from five previous publications, augmented by 20 interviews with *'experts'* (p.6) and anecdotal information from colleagues. It stands apart from other publications in this genre, by combining empirical research with comprehensive coverage of the background to, and practical conduct of, the examination, especially the viva. The book provides an overview of key elements, focusing on the viva in eight of the 13 chapters. The authors asked whether the viva is always an examination, as opposed to a ritual, and if it is the *'key site'* of decision-making (pp.29–30). They concluded that examiners regarded the viva as having three main purposes – examination, development and ritual, but that the predominant purposes in each are determined by: how examiners assess the thesis, their expectations of the candidate's knowledge; their personal examining style; their personal agenda; and the interpersonal dynamics present in the viva. Like Mullins and Kiley (2002), they suggested that inexperienced examiners may sometimes feel they have something to prove and therefore be harsh critics of the thesis.

How to examine a thesis (Pearce, 2005), the only one of the five publications aimed exclusively at examiners, offers advice on how to approach their role and avoid pitfalls. It begins with four viva case studies caricaturing different examiner approaches to the viva. Three of the scenarios describe situations that disadvantage the candidate, whereas the fourth presents a summary of a successful viva. The scenarios show the close relationship between the thesis evaluation and the examination of the candidate in the viva. The author suggests that although *'anything resembling "standard practice" remains little more than a fantasy...the fact that there is [none]...does not mean that there is no common practice: there is'* (Pearce, 2005:8). Pearce emphasises the importance of induction and training for doctoral examiners and highlights the variability in institutional regulations (Tinkler and Jackson, 2000).

3.5 The purposes of the modern PhD: can the examination assess multiple outcomes?

Until recently, it was tacitly understood that the PhD prepared future *'stewards of the discipline'* (Walker et al., 2008:12). The corresponding aim of the examination was therefore to provide evidence of the candidate's contribution to knowledge and independent research capability, and to assess whether their achievement was significant enough for admission into *'the academy'* (Delamont, Atkinson and Parry, 2000:1–9). It was also assumed, partly because many doctoral graduates entered the academic profession, that the examination was fulfilling its purpose. Since the 1980s, however, this assumption has changed, influenced by several factors, including decades of growth in candidate numbers arising from successive government policies to expand and democratise higher education (Morley et al., 2002; HEFCE, 2013). The ensuing diversity of candidates, purposes and employment destinations have led some, including candidates and graduates (Guccione and Bryan, 2017), to question the purpose of the modern PhD and its continuing utility as a qualification.

Two key and related issues threaten the ability of the PhD and its examination to satisfy stakeholders that they remain fit for purpose in the 21st century. The first is the narrowness of doctoral training (Nyquist, 2002; Borkowski, 2006; HEC, 2012) and the lack of focus on a professional career, exacerbated by the limited scope of the examination that is thought only to test the candidate's research capabilities rather than their potential in the workplace. The second is the perceived inability of the higher education sector internationally to define what constitutes *'doctorateness'* (Denicolo and Park, 2010; Wellington, 2013).

Preparation for employment

Now that fewer doctoral graduates proceed to academic roles (Vitae, 2010)¹², many argue that the PhD needs to be re-imagined as *'a labour market qualification'* (Park, 2007:17). Nyquist and Woodford (2000) pursued this theme in their survey of 365 practitioners, focus group participants, email questions and analysis of 400+ articles. Their publication arose from the *'Re-envisioning the PhD'* project¹³. Having found agreement that the doctorate prepared graduates effectively *'to conduct quality research'* (p.5), they suggest that *'individuals within and outside the academy claim that doctoral education inadequately prepares students for the other responsibilities and aspects of their careers'* (ibid), making a clear distinction between research and other attributes. They also discovered that *'an oversupply of PhDs for academic positions was an unintended consequence of responding to significant societal needs'* (ibid), citing three main causes: an increase in undergraduates necessitating more postgraduate researchers to teach courses; large-scale investment in research for scientific advancement leading to more postgraduate research assistants in some fields; and a *'commitment to improvement'* in higher education resulting in rankings that *'privilege the research model and drive a prestige economy'* (ibid). It is striking that, almost two decades later, this over-supply of PhDs now arguably exists in the UK and in other parts of Europe (Halse, 2007). Halse suggests that achieving *'the establishment of common, international quality standards'* (p.332) for the doctorate would be a significant challenge globally, suggesting the degree was undergoing *'radical transformation'* (ibid). Nyquist (2002) also pursues this theme, recognising that, although only universities are capable of awarding a doctorate, *'a PhD is the product of multiple owners or stakeholders, not the least of which are the doctoral students themselves'* (p.14). She suggests that although the PhD continues to be respected as a research qualification, this is no longer enough. Her summary of what PhD graduates need relates to current concerns in the UK:

The range of skills needed to function effectively today has increased enormously, especially for those who will occupy leadership positions. The new recipient of a PhD...in [any] job must move from a research and writing focus into a multidimensional range of activities and time commitments – committee work, team meetings, reports, teaching, worker training, planning and budgeting, recruiting and managerial oversight...Individuals within and outside the academy today contend that the doctoral experience should better prepare students for their professional destinations... (Nyquist, 2002:14)

¹² According to Vitae's website, after 3.5 years, only around 38% of 2008 and 2010 doctoral graduates were employed as teachers and/or researchers in higher education.

¹³ The Nyquist and Woodford publication(s) relate to the large-scale Carnegie Initiative on the Doctorate (CID) - that emerged in the US early in the 21st century (Walker et al., 2008).

Complementing this analysis, Borkowski (2006) observes that many stakeholders perceive *'a significant disparity between the training that doctoral students receive and the reality of the career options that await them'* (p.11), and remain concerned about the narrowness of the doctorate and its inability to prepare students for the job market.

Both authors are concerned with doctoral education in the US. If we accept that their analyses are pertinent in the UK, and given that UK doctoral programmes include training that focuses on the acquisition of professional and personal qualities, it is pertinent to ask if the functionality of the PhD examination is at fault by not overtly assessing such skills. Another factor is the often unacknowledged truth that completing a PhD or other doctorate necessitates possession of a range of personal attributes that are equally essential for succeeding in any kind of employment, including qualities of: time management; flexibility; leadership; tenacity; and problem-solving, among others. Inevitably, the doctoral graduate's discipline also affects the particular attributes they acquire.

Addressing the question of whether the examiners consider themselves to be assessing the work or the candidate, or if *'they are using the thesis as a sample of the student's potential or capacity for independent research'* Mullins and Kiley (2002:384) identify some disciplinary differences. Humanities respondents *'believed that "it is the text that is being examined, not so much what the students can argue and clarify in an oral"'*. Yet *'all of the mathematics/engineering interviewees and science interviewees'* agreed that where *'ambiguities or uncertainties'* existed *'it was useful to be able to clarify them with the student'* (ibid). The authors conclude that examiners believe they are assessing the candidate rather than *'the thesis document'*.

Lovitts' study on faculty criteria for evaluating dissertations recognises the continuing debate about what the PhD examination should be assessing:

Some faculty said they evaluated only the dissertation, some said they took the person into consideration. Should faculty's judgements be limited to the product (the dissertation) or is it OK to consider the person and the process? Along these lines, it was not uncommon for faculty to say that they held students to different standards based on their assessment of the students' capabilities and career goals. (Lovitts, 2007:117)

Her *'high-PhD-productive'* respondents claimed that some colleagues did not adjust their standards, also raising questions about differences among faculty.

Both Denicolo (2003) and Hall (2006) emphasise the importance of a face-to-face discussion in the viva given the dual purpose of the assessment – to evaluate the candidate’s contribution through the thesis and to assess their personal qualities as an independent researcher. Morley et al. (2002:272) extend this concept by suggesting that the examination’s fitness for purpose could be improved if *‘the “oral” element of assessment were to take on a more important function in relation to “skills” and “expertise” for a range of types of professional employment’*.

Morley et al.’s suggested approach begins to address the question of *‘doctorateness’* in any context, academic and other. Several authors seek evidence for this *‘elusive concept’* (Denicolo and Park, 2010:1).

Doctorateness

Questions concerning the nature of the doctorate and the attributes sought in the candidate are complex. They often relate to outcomes and how far there are shared standards and expectations for the award of a PhD. While authors have suggested that examiners should judge not only the candidate’s output, but also the doctoral process, they do not always make an overt connection between the candidate’s training and their attributes. The relationship between these elements and how they are currently assessed remains difficult to define.

The problem of whether the PhD meets current needs is therefore related to the question of whether the examination aims to assess the *‘product’* – the candidate’s thesis – or a) the [training] process (Denicolo, 2003; Lovitts, 2007; Park, 2007), or b) the person and their potential (Morley et al., 2002; Boud and Lee, 2009). Observing that, although the question *‘what is a PhD?’* might not seem difficult, it is *‘anything but straightforward to answer’*, Tinkler and Jackson (2004:7) suggest that definitions vary between a description of the PhD training process and the standard of what the candidate has produced. Kiley (2009) proposes that the doctorate’s purpose is *‘to both make an original contribution to knowledge and as a foundation in research training’* (p.35). Park (2007:37) relates doctorateness to the qualification itself, asking *‘what factors must be present for a particular degree to fit into the category of “doctorate”?’*, while also asking whether it is *‘really about the product (thesis) or the process (developing the researcher)’*. Lovat et al. (2008:67) argue that PhD examiners *‘are assessing the potential of the candidate in terms of their demonstration of scholarly qualities in relation to their research project, as well as the quality of the project and its outcomes’* by evaluating the person as well as their work.

Denicolo and Park (2010) contend that doctorateness can be demonstrated by both the doctoral degree and the candidate. They suggest that all doctoral awards and candidates should fulfil this requirement, emphasising that doctorateness is *'a quality rather than a state or tangible thing'* (p.2). They describe the *'essence'* (ibid) of doctorateness, as constituting the components that bestow a doctorate's identity. They suggest that the difficulties inherent in defining doctorateness arise from the complex set of personal qualities that might be found in a doctoral graduate, such as *'intellectual quality and confidence, independence of thinking, enthusiasm and commitment, and ability to adapt to changing circumstances and opportunities'* (ibid). On the other hand, they emphasise that it is the *'creation and extension of knowledge'* (ibid) required that sets the doctorate apart.

This is a critical point in the debate concerning the purposes of the PhD, as the creation of knowledge encapsulates potentially the most significant requirement of examiners: originality or contribution to the field. In adapting to the need to prepare graduates for multiple careers, including academia, it must retain its primary purpose as a research qualification to avoid becoming *'the next in a simple, additive progression'* (Denicolo and Park, 2010:2) of academic awards. The authors contend that the *'scholarly components'* (ibid) in Trafford and Leshem's (2008) model, are necessary but insufficient qualities on which to base judgements about an individual's research. Their assertion that *'the ...challenge is to reform doctoral assessment to meet contemporary situations whilst maintaining continuity and congruence with the past'* (ibid) epitomises the tension between assessing candidates' personal qualities while retaining a focus on their research achievements.

Wellington (2013) touches on the concept of doctorateness in the context of the doctorate's purpose and the reasons why candidates embark on their degree. He notes that there are many motives: intrinsic, relating to personal challenges for example; and extrinsic, including employment goals. Having summarised five potential purposes for undertaking a doctorate, Wellington suggests it is necessary to consider whether it is about the process, including the candidate's personal development and preparation for a career, or whether it is principally concerned with extending the *'body of knowledge'* (p.1493). He does not suggest both could have equal importance but contends that these questions are highly relevant to the doctoral examination process, which needs to accommodate the assessment of both product and process.

Literature exploring doctorateness suggests it can be assessed through three dimensions: person, product and [training] process. Given the variability of each candidate's experience

including disciplinary differences, however, and the summative judgement in the final examination, the arguments for assessing the person and product have more resonance. As we have seen, doctoral outcomes have attracted criticism in the US and the UK. This is arguably a covert attack on assessment methods, exacerbated by multiple thesis models that often reflect global expectations in a particular field (CGS, 2016; THE, 2015). Different groups and individuals have varied expectations of PhD graduates and their achievements so the criteria for the final assessment need to address a range of expected 'outputs' and 'outcomes'. The continuing high status of the PhD as an award arguably depends on the credibility of the assessment and examiners' ability to make an accurate judgement of the candidate's output(s). Additionally, it rests on the perception of the individual and the personal attributes they are able to demonstrate during and after the examination, whether or not these can collectively be described as 'doctorateness'.

4 Conclusion

Exploration of the fitness for purpose of the PhD examination must take account of the complex nature and purposes of the doctorate, which has evolved significantly. Literature reviewed in this chapter suggests that examiners' judgements and how they are made are central to the debate about the doctoral examination and its outcomes (Tinkler and Jackson, 2004:8). It also suggests that the candidate's perspective of their experience provides valuable insight and that examiner approaches to the viva are variable.

Questions concerning: whether it is possible to demonstrate some consistency in the final examination, e.g. in the process and/or the attributes sought by examiners; whether there are common standards and behaviours among examiners and graduates within and across different fields and institutions; and if the examination remains fit for purpose in a changing context where both the candidate and their work are being assessed, remain difficult to answer convincingly. In particular, the idea that both within and among subjects individual examiners might have similar interpretations of terms used to describe the achievements and skills a successful candidate needs to display, is contested.

The literature revealed that before this study no-one had attempted to research the examination as a two-part process, or to consider how examiners' judgements of candidates' achievements evolve by exploring the relationship between the primary assessment of the thesis and the candidate's performance in the viva. Morley et al. (2002) mention this explicitly:

The balance between the assessment of the text and the viva or oral examination is at least unclear and at best ambiguous. It is ill defined as to whether it is a one-or two-stage assessment process, and what the relative weight of the thesis is to the viva. (Morley et al., 2002:266)

This is a concept often neglected by other researchers: authors rarely comment on the relationship or the balance between the two elements of the final assessment. The literature also demonstrates a need for further research into how far examiners' judgements on the thesis can be altered by the viva.

Perhaps most importantly for the current study, there has been little or no independent observation of vivas in the UK. Trafford and Leshem (2008:201) note that the viva is '*under-researched*', while Lovat et al. (2015:7) contend that '*unresolved issues and gaps*' in knowledge about the doctoral examination still exist '*not least in relation to the role and robustness of the Viva*' (ibid). However, it was Hartley and Jory's (2000:88) observation that '*it would be interesting, but probably difficult, to obtain some observational data from independent observers about what goes on in viva situations*' that had a major influence on my research design. Additional evidence was needed to test some of the anecdotal evidence mentioned by Hartley and Jory, especially since the results of empirical studies to date remain ambivalent in relation to fairness of outcomes. The next chapter addresses my theoretical approach and how the research aims were operationalised in the study.

Chapter 4: Theoretical approach, research design and methods

1 Introduction

Investigating the complicated social phenomenon of the PhD examination demanded careful choice of methodology and methods. The research design combines independent observation of the viva with semi-structured interviews to enable in-depth study of 10 'bounded' cases, as described in Figures 4.2 (p.95) and 4.3 (p.98) below. The observation of candidates' vivas posed a series of challenges. The case studies were augmented by interviews with experienced examiners and a focus group discussion with five academic staff from different subjects in one university, unrelated to the cases. The private nature of the final PhD examination in the UK led me to focus on the viva as the entry point to wider scrutiny of the two-part assessment. Two-page summaries of the cases, with an accompanying analytical framework, are at Appendix 1.

2 Theoretical perspective

I aimed to explore in depth a number of individual cases rather than to seek generalisable results. After consideration of several research strategies, a realist theoretical perspective combined with a multiple case study approach was chosen. The objective was to generate sufficient qualitative data from interviews and observations to provide an in-depth study of each case.

The contested nature of research methodology in the social sciences is well documented (e.g. Lagemann, 2002; Slavin 2002; Kaestle, 1993) and perhaps unsurprising given that the social sciences and education in particular span a range of fields that use both qualitative and quantitative data. According to Pring (2004:37), however, *'the rather rigid separation of the quantitative from the qualitative approaches to research'* by those who try to maintain too great a distinction between subjective (individual consciousness) and objective (measurable) paradigms misses the *'subtle interconnection'* between them, often excluding one completely in favour of the other. Miles and Huberman (1994) suggest that links between qualitative and quantitative data can help to corroborate or triangulate data and provide new ideas.

In response to criticisms of research in the social sciences and of education research in particular (Scott and Morrison, 2005), some researchers have developed approaches such as realism that combine rigour and reliability with the flexibility to use methods that take account

of the context in which the research is occurring, so as to explain sophisticated nuances in social phenomena that are best interpreted through observation in cultural contexts (e.g. Blaikie, 1993; Guba and Lincoln, 1989). It is now widely recognised by followers of different research paradigms that both hypothesis testing using quantitative data and the generation of qualitative data to answer cogent research questions can be relevant in education research, because of the heterogeneity of research topics, ranging from large-scale quantitative enquiry to individual case studies (Yin, 2009).

The complexity of the research topic and the sensitivities involved in gaining access and conducting the study had a significant impact on the choice of theoretical perspective. It might have been possible to take a constructivist approach (Crotty, 1998; Guba and Lincoln, 1994) and to develop theories about the final PhD examination based solely on the views expressed by the actors involved. However, a realist perspective – taking account of what we know to be reality irrespective of any personal opinions or beliefs while recognising the impossibility of complete objectivity – enabled me to combine the views from the viva participant interviews with an emphasis on my independent perspective as an observer who had never been actively involved in the two-part PhD examination, yet had theoretical knowledge.

In order not to confuse the realism he is arguing for in educational research with '*naïve realism*', Pring (2004:61) advises avoiding the presumption that reality can be described completely accurately. Instead he advocates a '*critical*' realism that enables differentiation between objective and subjective approaches (Bhaskar, 1986). Objective enquiry in this context is defined as research that reflects reality and is achieved through, for example, ensuring that research outcomes are based on evidence provided by the data, testing emerging conclusions against experience, subjecting the research to external scrutiny, and making sure that it is reported coherently and does not contain contradictions. Pring (2004) also argues that a realist perspective forces us to accept reality as something independent that constrains and limits the researcher, thereby imposing a discipline that requires our conclusions to be justified by reference to the evidence provided. He emphasises this does not necessarily place the researcher in the '*positivist*' camp, rather that the concept of realism can be applied in both quantitative and qualitative approaches. This concept of critical realism is further developed when Pring (2014, 2015) suggests that education researchers can find a '*middle way*' avoiding the '*false dualism*' that potentially compromises some educational research and '*to preserve a sense of "realism" which is compatible...with constant uncertainty*'

(Pring, 2014:1). Adopting a 'critical realist' perspective enables a specific research problem to be investigated as objectively as possible, taking account of the nature of the problem.

The current study draws on Pring's interpretation of critical realism and also links to the orientation of transcendental realism (Bhaskar, 1978/2008 and 1989), adopted by Miles and Huberman (1994). Using this concept, the researcher takes the view that a social phenomenon, in this case the PhD examination, can be regarded both objectively and subjectively, that comparisons may be made between different occurrences of the phenomenon, and that some common characteristics can be found within and among them. According to Henwood (1996), this interpretation of transcendental realism enables researchers to defend the reliability of qualitative research without compromising on priorities such as analytical approach, the ability to be sensitive to different contexts and the perceived credibility of the research by others in the field.

In this study, I aimed to demonstrate integrity in the research design and implementation, and awareness of the diverse contexts and disciplines in which the final examination occurs. The examination is treated as a social phenomenon (Hammersley and Gomm, 2000) that is studied in depth through a number of 'bounded' cases across different subjects, including exploration of similarities and differences identified by participants and through observation (Miles and Huberman, 1994).

3 Qualitative data generation

The challenges for research using qualitative data are well documented and include the risk of researcher bias and the potential for inadequate sampling and data overload (Guba and Lincoln, 1989; Miles and Huberman, 1994; Stake, 1995). The juxtaposition of qualitative data generation with a realist perspective could be perceived as an additional challenge in this study, yet here critical realism enabled potentially subjective knowledge to be combined with objective enquiry leading to research outcomes that go some way to reflecting what is actually happening (Pring, 2004). This supports Slavin's (2002) assertion that qualitative data can be valuable and achieve credible results in contexts where experiments cannot be used to answer research questions. In this way, incremental advances in research may contribute to evidence-based policy, thereby improving practice.

Qualitative data must provide enough data or evidence, including clear details about how the research has been conducted (Lincoln and Guba, 1985), to enable the reader to evaluate the researcher's analysis and interpretation and be presented using logical arguments. I aim to provide *'a coherent and illuminating description of and perspective on a situation that is based on and consistent with detailed study of that situation'* (Ward Schofield, in Gomm et al., 2000:71).

The purpose of the current study was to analyse the PhD examination and to draw some conclusions about the process through which examiners judge candidates' achievements using the thesis and viva. The *'strong potential for revealing complexity'* (Miles and Huberman, 1994:10 and see Hammersley, 2010) inherent in qualitative data enabled me to develop a detailed profile of the final PhD examination as seen from different individual and subject perspectives. Qualitative information enables researchers to capture *'real life'* situations, of particular relevance to this study with its emphasis on specific cases, each being *'a focused and bounded phenomenon embedded in its context'* (ibid).

4 Case study methodology

Case study is thought by some to be one of the most challenging ways of carrying out social sciences research (Yin, 2009), partly because of the difficulties of demonstrating rigour and validity in the way the research has been conducted and the potentially challenging data analysis. The value of case study, however, lies in its ability to generate rich data that enables the researcher to explain or interpret a phenomenon that does not lend itself to being investigated by quantitative data generated in large-scale statistical studies (Stake 1995). If used systematically, case studies can be effective for learning about human interaction and to probe a concept or phenomenon in depth. Exploration of multiple cases within a realist paradigm therefore proved a highly appropriate methodology for investigating the doctoral examination.

4.1 How and why a multiple case study approach was used in this research

Case study has been defined as *'research that investigates a few cases, often just one, in considerable depth'* (Hammersley and Gomm, 2000:3). According to Scott and Morrison (2005:17), *'a hallmark of case study'* is that cases emerge from *'naturally occurring situations in which variables are not, or cannot be controlled'*. An important criterion is that *'cases are*

not created "artificially" for the purposes of the research, as with experiments', but are 'naturally occurring' (Hammersley and Gomm, 2000:4). Case study may be treated either as a research method or 'as a distinct research paradigm' (ibid). Yin (2009), for example, refers unambiguously to case study as a method, as does Crotty (1998). Dowling and Brown (2010) take a similar position, suggesting there is no such thing as case study methodology. They argue that the transformative intervention of the researcher means that using the term 'case' is merely a way of describing sampling procedures rather than the researcher being able to claim any kind of objective study of a phenomenon.

The extent to which these differing perspectives of case study apply varies according to individual perspectives. Here, case study is used as a research *methodology*, the research *methods* being interviews and viva observations. This combination of case study methodology and carefully selected research methods was chosen to generate data about the phenomenon of the doctoral examination (Hammersley and Gomm, 2000).

Critics of case study research (e.g. Atkinson and Delamont, 1985; Bassegy, 1999; Hammersley and Gomm, 2000; Scott and Morrison, 2005; Flyvbjerg, 2006: 3-5) suggest that: (i) it lacks rigour and authenticity; (ii) it cannot be used to extrapolate findings to other cases or situations; and (iii) it therefore is not useful for developing theories or testing hypotheses in a similar way to large-scale quantitative studies. Bryman (2008) suggests that the extent to which case study methodology is criticised for its ability to meet criteria of reliability, replicability and validity depends on the perspective of the researcher, so that authors on methodology who are principally concerned with qualitative data are less critical of the ability of case study to demonstrate these attributes than those who are more familiar with research involving quantitative data. In this study I aimed to address these criticisms through rigorous application of research methods, by avoiding hypothesis-testing, and by not attempting to generalise or claim that the cases have been studied completely objectively, especially given the limited number of cases and the personal perspectives of the actors and the researcher.

More positively, comparing cases can give a researcher 'a nuanced view of reality' (Flyvbjerg, 2006: 8-9), with each case generating qualitative data about a phenomenon that adds to existing knowledge. Gomm et al. (2000), for example, argue that, while investigating multiple, systematically selected cases can be difficult, subsequent studies can build on earlier work and increase the number to a sample over time to allow generalisation to be possible. They also

suggest that this is a rare occurrence, either because subsequent research does not necessarily complement earlier studies or because data that would enable meta-analysis is missing. In this study I aimed to create a design that could be replicated by others, potentially generating a larger sample of cases for analysis.

4.2 Bounded cases

The main part of the study is based on ten 'bounded' cases. Each consists of a self-contained example of a PhD examination involving a candidate and their examiners. Some cases include other actors who are permitted to attend but not participate in the examination, for example, supervisors. Each candidate submitted a thesis and was undergoing an oral examination in which they defended their thesis in response to questioning by the examiners. Each case was a naturally occurring event where the final examination was the culmination of a PhD programme.

The question of the 'bounded' case in this study requires both exploration and explanation, using the abstract definition of a case as a phenomenon that occurs in a bounded setting. According to Miles and Huberman (1994:25), each case has a '*heart*' and an '*indeterminate boundary*' that defines what is included and excluded. In this study the boundaries are determined by occurrences of a phenomenon - the final PhD examination - with the oral examination being the entry point to accessing each case. At the 'heart' of the cases are the candidate and the examiners; the examination takes place within a room, for a period of time that varies depending on the individual circumstances of each viva. Within the boundary of the examination, other 'peripheral' actors (supervisor, independent chair/ convenor, observer) may be present. Other factors, physical and ephemeral, are present in each case such as the candidate's thesis, which is examined during the viva, the interaction between the examiners and the candidate, the notes taken by the observer during the viva and the final outcome of the examination. Figure 4.1 shows the 'heart', components and outer boundaries of the case.

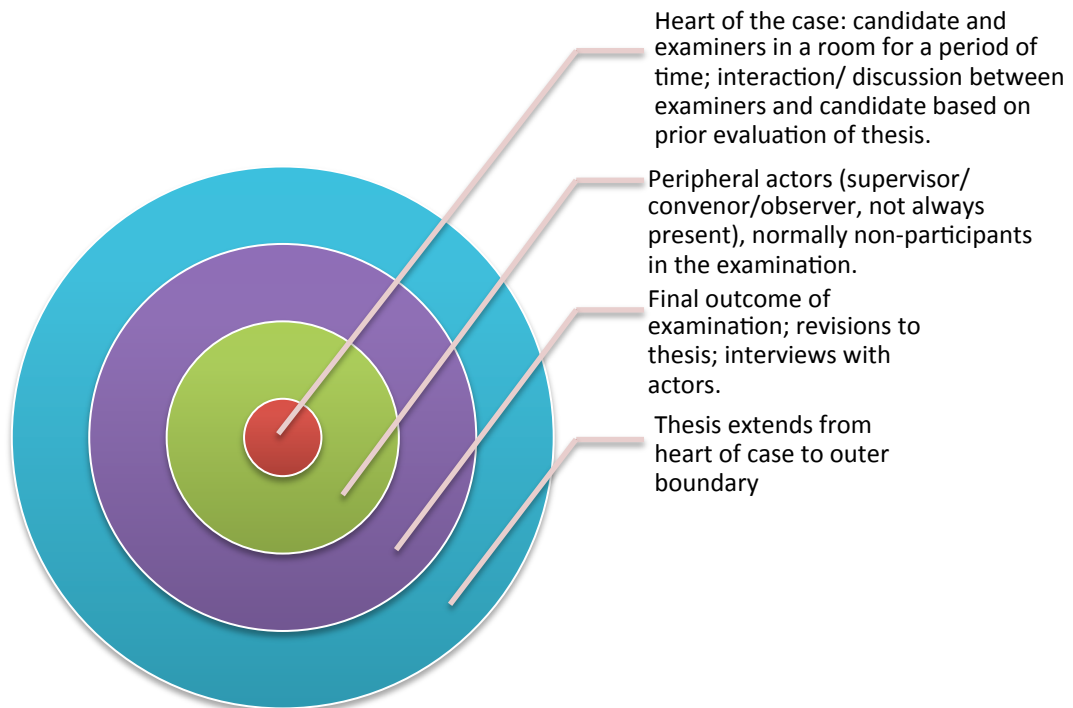


Figure 4.1: The `heart', components and boundaries of each case

The thesis is at the heart of the case as well as extending to the outer boundary: it is present throughout the case and beyond. Each case is situated in an institutional and subject-specific context: the hinterland that influences both process and judgements.

The research design included interviews with all actors in each case. However, some individuals declined while others were not interviewed for practical reasons. In one or two instances after some time had elapsed since the viva I decided that it was not worth the investment in time to pursue an interview, especially since I had more than enough data.

In addition to data generated by observations of different vivas, the examiners', candidates' and other actors' `voices' are heard through analysis of the interview responses, thereby aligning with another of Scott and Morrison's (2005) requisite characteristics, which is to give those involved in the case a `voice'. In this study my `voice' is also heard through the viva observations. This dual approach adds another dimension by balancing different perspectives, with the intention of applying Pring's theory of objective enquiry that is relevant in a realist theoretical framework. The cases observed in this study are `*complex, functioning*' units which have been investigated in their `*natural context*' with more than one method. They are also `*contemporary*' in that they represent current practice in examining the PhD in the UK (Johansson, 2003: 2).

Studying ten cases provided breadth as well as depth of data. In this research, complexity of the information was also a factor, although not necessarily an advantage. The value of researching multiple cases (Pring, 2004; Stake, 2006) is demonstrated by this study, where the individual characteristics of the final PhD examination are explored. The *'quintain'* – the whole study or *'collective target'* containing multiple cases (Stake, 2006:6) – enables a researcher to gain a more comprehensive understanding of different elements of the phenomenon as a result of analysing multiple cases (Bryman, 2008). For example, the cases overall provided me with an insight into the influence of experience, discipline and background on different examiners' approaches to doctoral assessment. The institution in which the final PhD examination is taking place also contributes to shaping the doctoral examination process (Stake, 2006) and therefore the format of the PhD final examination in this study is affected by institutional guidelines and conventions, as well as subject-level expectations and doctoral traditions.

5 Dimensions of the study

This was a small, in-depth, exploratory study. Two dimensions were important in the design:

Subjects to be studied: The cases spanned different subjects, in order to explore possible differences in the final PhD examination process in different disciplines and to explore whether examiner expectations across subjects are similar or different. Gomm, Hammersley and Foster (2000:106) recommended systematic selection of cases as a complementary approach for case study researchers when addressing *'the problem of empirical generalization'*. However, in this study the challenge of gaining access to observations resulted in opportunistic sampling. The cases have been categorised as either a) science, technology, engineering and mathematics (STEM), or b) arts, humanities and social sciences (AHSS). These broad subject groupings have also been used in the analysis. Although in multiple case studies it is not always necessary to investigate each case in equal detail, in this study all cases were analysed in equal depth to optimise the data.

University: Six UK universities participated in the cases - one was post-1992, five were pre-1992. Only one of the five universities that originally agreed to the study provided access to the cases. Another of the original five provided a focus group; the remaining three proved unable to participate due to challenges concerning access, and a further university allowed me

to undertake two pilot observations that subsequently contributed to the ten cases. A total of 11 universities were involved in the study, when the experienced examiners' institutions were taken into account. Sampling the PhD examination at more than one university was useful, especially when exploring the process, although subject specificity was more influential in some factors, for example, the length of the viva.

6 Ethical considerations

My two-stage application for ethical consent to Oxford's Central University Research Ethics Committee (CUREC) was rigorously scrutinised. This was beneficial for the research design as a whole as it led to careful consideration of how the various research methods would be used, in particular, how participants' identities would be protected and to ensure that no-one would be coerced into accepting the request to have an observer present at their viva.

In assuring the universities and candidates that, if allowed access, their confidentiality would be respected and the observations conducted in an ethical and professional manner, I was supported by the rigour of the CUREC's approach in assessing my applications for ethical clearance for DPhil studies.

Ethical requirements meant that observations could not take place unless all the actors involved in a viva agreed to be observed: the candidate, all examiners, supervisor(s) and independent chair or convenor, if present. The universities involved agreed to contact viva participants independently on my behalf so I did not know who was being approached and when. This was appropriate in the sense that my not having direct access to anyone involved meant that those being asked to consent to the observations did not feel that the researcher was putting pressure on them to agree. The participant information sheet and consent form are at Appendix 2.

Questions such as the need to maintain anonymity of doctoral candidates/graduates and the privacy that is normally expected in the oral examination, were all taken into account in designing this study. All participants could opt not to consent to an observer being present and the anonymity of individuals was guaranteed.

7 Access challenges

The question of access is often linked to ethical considerations and associated with hierarchies, power and authority (Scott and Morrison, 2005). The first step in gaining access to viva observations was therefore to acquire permission from senior university managers to conduct research in their institutions. This was assisted by my professional background, including work on behalf of national organisations in the higher education sector, which had provided opportunities to develop a rapport with a large network of academic and administrative contacts in a range of universities. I was careful not to exploit these relationships, and on commencing the research had no formal affiliation with these institutions. Instead I attempted to demonstrate awareness and sensitivity in an exploratory letter (Appendix 3), sent to senior staff (e.g. at pro vice-chancellor level) at six universities in June 2011 explaining the proposed study and requesting the opportunity to contact academic staff such as doctoral supervisors and examiners, including those with faculty or university level responsibility for doctoral education.

Five of the six universities responded positively, several of them reassured by the ethical approval process undertaken with my university. The possibility of one or two viva observations and associated interviews was proposed at each institution, a step further than anticipated as the exploratory letter had only suggested using a questionnaire. The sensitivity of such a request was significant given the privacy of doctoral vivas. However, the universities accepted my final design and I was put in touch with staff at each institution who could facilitate access and arrange for candidates, examiners, supervisors and independent chairs to receive the necessary information.

One of my supervisors arranged access to two viva observations unrelated to the institutions already approached that would enable a pilot study to be conducted. These took place in November and December 2011. The pilot viva observations and interviews (which contributed to the overall number of ten cases), were useful for testing the research design and instruments and led to small changes in interview questions. Even in these two cases, it was not possible to interview all participants, yet arranging post-viva interviews proved to be the least of the challenges encountered in operationalising the study. The pilot was helpful in testing the research methods, but did not prepare me for some of the frustrations I would experience in gaining access to viva observations.

7.1 Operationalisation

In a study of this sensitivity it was inevitable that operationalisation would be a challenge. The effort needed to negotiate ten viva observations and related interviews was beyond expectations, especially the time and persistence required to complete the data collection. Having undertaken the preparatory work, I was optimistic that the observations would be realised at the five universities that had agreed to take part. It soon became clear, however, that agreement from senior staff was only the first step and that gaining the permission of all viva participants to allow an independent observer to be present would prove the biggest challenge in implementing the research design. Rather than taking 6-12 months as had been expected, the data collection period lasted from November 2011 until May 2014.

I intended to undertake 15 case studies involving observations of vivas relating to the PhD assessment, interviews with all the actors in each viva about the final assessment overall (thesis and viva) and to request sight of examiners' joint final reports relating to each candidate. However, these were ambitious aims and despite extensive efforts it was not possible to implement them in full. In addition to the difficulty in gaining access to vivas, I was unable to access examiners' post-viva reports and not all participants were interviewed in each case.

Of the five universities that had initially agreed, four presented considerable challenges. For example, one institution, having been encouraging in the expectation that viva observations would be possible, told me after some time that none of the observations could take place because two whole cohorts of students (in different subjects) had declined to be observed. In another, one supervisor agreed to ask a student whose viva was imminent if I could observe, but the candidate declined and the supervisor emailed me to confirm: 'I'm sorry to say that the candidate for the upcoming PhD viva I mentioned has finally decided that s/he'd rather not participate'. Although not able to provide access to viva observations, this university allowed me to visit and have a focus group discussion using very similar questions to those asked of individual examiners and supervisors in other institutions post-viva, which elicited some interesting perspectives.

Two further institutions, despite sustained efforts by senior staff to negotiate access with different academic departments, were unable to fulfil the commitment to provide

observations, as a result of either staff or candidates declining to have an external observer present. Examples of feedback received from one of these institutions included:

On behalf of a candidate's two supervisors: 'With thanks to Gill for the invitation, on this occasion we will decline. I do wish Gill well with her research.'

From a candidate: 'Unfortunately I cannot agree to participate in Gill's study. It's a too delicate balance for me and while I note that Gill will take care to be quiet I do not want any distractions on the day. I am sorry Gill. I hope you have more luck with others.'

A senior contact at one of these institutions, who was supervising a candidate at another university, was instead able to arrange for me to observe the supervisee's viva, with all participants' permission.

The approach taken by the fifth university was to manage the process very carefully, including providing extensive administrative support (including explaining the project to candidates) and submitting all the details of the study to their own university research ethics committee. This approach appeared to have a positive influence on access. But even at this institution, it was not possible to persuade all the potential viva participants to help with the study: around ten possible observations were ruled out, either because the candidate or one of the examiners declined the request.

Examples of extracts from emails received from the administrator organising the observations:

'Unfortunately [the candidate] and one of [their] examiners have replied to decline their consent for the observation/research; the examiner was positive about the research but was concerned about additional strain on the student.'

'I am so sorry but unfortunately [the candidate] replied to decline the request as s/he said "*I'm afraid I would be very uncomfortable having an observer at my viva, so if it OK I would prefer not to have Gill present. Sorry I can't be more helpful*". I know this will be disappointing.'

'Unfortunately [the candidate's] external examiner has replied to decline the observation saying: "*With regard to the observation of the viva, I think this is an important study to carry out but due to the timing of this exam I would prefer to pass on this occasion and not have an observer*".'

Another setback was that in one case, where all permissions had been given, I could not attend the viva (in January 2013) because of a heavy snowfall.

During the academic year 2012-13 it was clear that only one of the five original institutions could facilitate observations, so I asked senior colleagues I had known for many years through

professional links, if they could provide access to observations. Through this route, I was able to observe three more vivas, while continuing to work with the remaining university of the original five, securing a total of four observations there over a period of 18 months.

As the field work period progressed, the requirement for all actors in the viva to agree to the observation proved a major challenge in gaining access. As the above quotations demonstrate, nervous candidates declined to be observed and in one or two cases when I thought the observations had been arranged, the students changed their mind as the viva approached. In some cases, I wondered whether the candidate's agreement might encourage examiners who were reluctant to agree to the observation to be better disposed towards the prospect of an independent researcher being present at the viva. Some examiners declined an observation because they were concerned about the additional pressure on the candidate. Also, securing the candidate's permission would not have helped to persuade those examiners who agreed in principle to an observation but would not confirm until they had read the candidate's thesis and who thereafter declined the observation (one or two examiners). In one example of 'in principle' agreement by the examiners (the case where eventually all gave permission for me to attend but that I missed because of snow), the examiners delayed giving final permission until the day before the viva. At this point I understood that the external's report had not been received and that the internal examiner wished to read this before committing to the observation as their own report was not 'overly positive'. The internal was concerned for the candidate if the viva did not go 'as... hoped'.

In all cases, once one of the viva participants had declined permission, it was not possible to proceed as everyone had to agree before the observation could proceed. The opportunistic sampling approach eventually ensured that ten cases were observed. This was a one third reduction of the original number I had hoped for, yet the ten cases generated sufficient rich and comprehensive data for the study.

7.2 Arranging post-viva interviews

Of the 43 interviews, 23 were conducted face to face, 12 using Skype, and 8 were completed by telephone. All but the telephone interviews were recorded. As anticipated, some post-viva interviews proved difficult to arrange; the candidates were the most accessible group.

Examiners were also co-operative and all except three of the externals were interviewed.

Most interviews took place within a month of the viva; a few happened quite a while

afterwards. Irrespective of timing, all interviewees appeared to have a distinct recollection of the examination. In one or two cases, because of being generally busy with work or other observations, it took me some time to follow up interviewees who for some reason had asked immediately after the viva to delay their interviews. Towards the end of the data collection period I agreed with my supervisor, that I would not pursue two of the interviews in order not to delay data analysis, especially since I already had 43 interview transcripts. The number and distribution of the interviews is provided in Table 4.5.

Extracts from emails received in response to requests for interviews include the following:

From an external examiner: 'I feel a bit overwhelmed with work at the moment - and just need a break...I don't think I can face an hour's interview right now - it's nothing to do with you personally, just a cumulative thing of getting a longer list of things that I just can't manage. If you can't find any other respondents do get back to me, but for the time being I think I'm going to decline. I hope you got enough data from observing the viva itself and I wish you all the best with the PhD'.

From a supervisor: 'I am terribly sorry but I am extremely pressed for time at the moment and cannot spare an hour. Also, the viva you refer to is still my only viva experience. I have to say no on this occasion.'

Neither the external examiner nor the supervisor above was interviewed. Their responses suggest both were under considerable pressure of work; and that in one case, inexperience led to unwillingness to be involved. As this section reveals, access to data was one of the most significant challenges in this study, yet significantly, most participants were positive about the research, two examiners who declined to participate nevertheless confirming their support.

8 Sampling

Initially I intended to explore differences between institutions and subjects, and to undertake purposive sampling (Bryman, 2008). My Transfer of Status document outlined a rationale for sampling a group of pre- and post-1992 universities across three UK countries, including several English regions. I had approached several institutions at this stage (three of whom had agreed to participate in the study), but had not yet encountered the problems of access described. I planned to sample a range of subjects at the selected institutions and was proposing to use a questionnaire and to interview some staff at each institution, as well as observing vivas. Following the advice of my Transfer examiners and supervisor I decided to remove the questionnaire as a method and concentrate on a smaller number of institutions with the aim of observing 15 cases. This meant that when the challenges of access arose, there

was more flexibility in the research design and I could use opportunistic sampling without compromising the project’s rigour or validity. It would never have been possible to select ‘typical’ cases, because of the idiographic nature of research degrees. To some extent, particularly at two institutions, I was able to select which PhD finalists to observe, thereby achieving some subject coverage, but these preferences were dependent on the approval of all actors. Clearly the ten cases, while covering different disciplines, are not representative. Subject coverage was extended through interviewing non-case related examiners.

Candidates were varied in age. I asked for their age within a 5-point scale: 25 or under; 26–30; 31–40; 41–50; 51 or over. At the time they were interviewed, two were in the 26–30 category, four were between 31 and 40, two were between 41 and 50 and one was in the 51 or over category. The HESA definition of mature candidates in postgraduate study is aged over 25 on entry. The more mature candidates were in social sciences, whereas younger candidates were in science, as would be expected given PhD candidate profiles generally. Data obtained from the Higher Education Statistics Agency (HESA) for the purpose of comparison with my contributors show that in 2014–15, for both full-and part-time candidates, those studying STEM subjects had a younger age profile than those in other disciplines. All my contributing candidates were therefore classified as mature, although it is interesting that the HESA figures show that over 70% of all doctoral candidates are also over 25, with around 36% being between the ages of 31 and 50. The age and discipline profile of candidates involved in my study compares with the HESA figures as summarised in Table 4.1.

Ages	HESA 2014-15					This study				
	STEM		Non-STEM		Total	STEM		Non-STEM		Total
	No.	% age	No.	% age		No.	% age	No.	% age	
25 or under	23,177	81%	5573	19%	28,750	0		0		0
26-30	18,852	65%	10,218	35%	29,070	2	100%	0		2
31-40	14,198	57%	10,807	43%	25,005	3	75%	1	25%	4
41-50	4,315	40%	6555	60%	10,870	0		2	100%	2
51 or over	1,807	28%	4,738	72%	6,545	0		1	100%	1
Age unknown					5					
Subject unknown					10					
Totals	62,349		37,891		100,255	5		4		9

Table 4.1: Age and subject profiles of PhD candidates: comparison with UK figures

The nature of the research design and the ethical requirements of the study prevented me from sampling cases where examiners were concerned about the quality of the thesis, potentially skewing the data towards higher quality candidates. As illustrated above, for various reasons, not all candidates or examiners were confident enough to be scrutinised by an observer, which potentially increased the bias inherent in the study.

9 Research design

The aim of the research was to address criticisms of the PhD examination, both anecdotal and in the literature by investigating how PhD examiners make judgements. I decided to focus on the potential for inconsistency and lack of attention to the candidate's personal attributes by exploring the characteristics examiners were expecting to discover in the thesis and in the candidate. This led to an analysis of the process and the relationship between the thesis and viva. The design was chosen for its ability to generate rich data on each case, therefore observing vivas and conducting interviews with the main actors involved in the final examination were chosen as the core methods. Interviews were conducted with all but one of the candidates and where possible the external and internal examiners; they also included the candidate's supervisor(s) and/or convenor or independent chair where present at the viva, if they were willing.

Case observations and interviews were complemented by interviews with seven experienced examiners unrelated to the case studies, who could look longitudinally at their composite experience as an internal and external doctoral examiner (and often also as a supervisor) in a single field, at different institutions. In addition, a focus group comprising five academic staff from different subjects was convened at one university, their discussion informed by the interview schedule used for examiners and supervisors in the cases.

Three reasons made it important to include some interviews with experienced examiners and a focus group discussion, unrelated to the cases:

- Even though six universities agreed to participate, it was uncertain whether enough observations would be forthcoming;
- The number of examiners who would agree to be interviewed in the various cases was unknown, so it was an advantage to be able to consult experienced examiners, especially in subjects not covered by the cases;

- The focus group discussion offered the opportunity to balance and compare opinions through real time discussions.

Interview questions were constructed to provide insight into: the process as a whole; the respective roles of the thesis and viva in demonstrating candidates' achievements and the relationship between them; the evidence and criteria used by examiners when assessing candidates; and the factors affecting the final examination outcome. The underlying objective of the research was to test the rigour and consistency of examiners' judgements and the fitness for purpose of the examination through the insights of the actors involved.

9.1 Validity and rigour

A realist theoretical perspective was integral to maintaining rigour and validity in the research design: comparing the researcher's observations with data generated by the interviewees in each case allowed the observations to be compared with the interviewees' responses, and vice versa. Use of multiple cases provided some opportunities for comparison of the evidence provided by the interview and observation data.

In the design I aimed to meet the criteria for an educational case study as defined by Scott and Morrison (2005:19–20) and to conduct an empirical study *'in such a way that sufficient data are collected for the researcher to'* explore significant features of the cases, create plausible interpretations and construct worthwhile arguments.

Rigour was maintained through theoretical and practical strategies such as: opportunistic sampling of the examination at different universities and in different subjects; collecting sufficient data for a multiple case study approach using several research methods; consistent use of the questions in semi-structured interview schedules (while not being constrained by them); recording (where possible) and transcribing the interviews and coding the data; carefully analysing the data; and avoiding claims of generalisability on the basis of relatively few cases. Simultaneously, the research design provided a flexible approach, which proved to be crucial to successful implementation.

The individuality of the PhD degree, the subject-specific expectations about doctoral graduates' achievements and abilities and differing university regulations and guidelines for doctoral degrees all need to be acknowledged in framing this study and analysing the outcomes. The factors that vary in each occurrence of a doctoral examination include: the

candidate; the examiners; the candidate's supervisor(s) and the independent chair, if present. The candidate's thesis or other output is, by definition, unique because it constitutes a contribution to the field. If we define the doctoral examination as a phenomenon and each occurrence as a 'case', then all these components are attributes of each case (Bryman, 2008:33).

The next section summarises the rationale for choosing those methods and describes how they were employed.

10 Research methods

Idiographic research methods (Bryman, 2008; Wolcott 1992), concentrating on individual cases where a phenomenon is studied over a period of time, were most appropriate for this research. Applying this approach using observations and interviews enabled me to gain an in-depth knowledge of the characteristics of the PhD examination from the perspective of different actors by engaging with individuals, during the viva and through the semi-structured interviews. By observing examiners' and candidates' interactions in the viva, then conducting post-viva interviews, qualitative data were generated using explicit and systematic methods. 'Realist' strategies in this study (Miles and Huberman 1994:6) include: nonreactive (unobtrusive) research [silently observing the viva]; human ethology [behaviour and interactions during the viva]; and interviews [with participants in the examination].

Observations and interviews were fundamental to understanding the final PhD examination process in this study. In particular, observation through being present is considered critical to effective and rigorous case study research (Pring, 2004). Interview questions arising from observations, such as how a viva had influenced the overall outcome of the final examination, enabled me to follow up points of interest emerging from the vivas and to probe more deeply examiners' judgements and the qualities they seek in candidates. For example, interviews enabled me to pursue a line of questioning arising from the examination of the candidate during the viva, or the way the viva had been conducted. In addition to the viva observations and interviews (including those with experienced examiners unrelated to the cases and the focus group described earlier), I had access to university regulations and guidance for the examination. The use of additional data sources added flexibility to the research design while increasing the range of views surveyed. Table 4.2 summarises the purpose of the research methods in addressing the research questions.

Main methods	Purpose
Observation of the oral examination (the <i>viva voce</i>)	<p>The purpose of the observations was to provide insight into:</p> <ul style="list-style-type: none"> • the respective purposes of the thesis and the viva • how examiners use the viva to clarify elements of the thesis and the candidate's thinking about methodology, and to come to a final view about the examination outcome • any difference of approach demonstrated by disciplines • the candidate's ability to respond adequately to the examiners' questions and how their performance in the viva affected the final outcome
Semi-structured interviews with examiners and supervisors	<p>The purpose of interviewing examiners and supervisors (the latter are often examiners themselves and may have observed multiple vivas) was:</p> <ul style="list-style-type: none"> • to explore examiners' and supervisors' views about the examination process, including its strengths and weaknesses • to gain an understanding of how examiners judge a candidate's achievements and of the attributes sought • to compare respondents' views with what I had observed in the vivas • to gain a broader picture of PhD examining conventions in disciplinary fields • to discover the extent of support among examiners and supervisors for retaining a threshold judgement on the award of the PhD.
Semi-structured interviews with PhD candidates	<p>The purpose of candidate interviews was:</p> <ul style="list-style-type: none"> • to obtain candidates' views about the final PhD assessment process, including its strengths and weaknesses • to assess whether candidates had an understanding of the process, the purpose of the examiners' questioning and the qualities examiners were seeking in order to recommend the award of a PhD • to discover the extent of support among candidates for retaining a threshold judgement on the award of the PhD
Focus group discussion involving individuals from different subject areas	<p>The purpose of the focus group discussion was:</p> <ul style="list-style-type: none"> • to add to the evidence about what examiners are seeking in PhD candidates • to offer an additional institutional perspective using an alternative format • to provide a situation where, if relevant, the significance of different subject perspectives on the final PhD assessment could be evaluated
Additional data	
Analysis of <i>institutional guidelines and regulations</i> enabled comparison of different institutional requirements for the PhD examination and its outcomes.	

Table 4.2: Research methods and their purposes

The choice of observations and interviews as the two main methods was valuable in the context of a realist theoretical perspective: observations were undertaken by me as an

independent researcher who took no part in the examination but as objectively as possible recorded and compared occurrences of this phenomenon. The positioning of different actors during the examination (Figure 4.3) was noted, as well as details such as how much the viva appeared to influence the final outcome of the examination. The interviews reflected the potentially subjective views of the actors in the research process but also highlighted any variations in approach emanating from the disciplines. In analysing the views of the different actors engaged in the final PhD examination, case by case, I aimed to give each actor a voice and to enable them to provide their own interpretation of the event, as well as to identify and explore similarities and differences.

11 The cases: sequence and procedures

The cases constitute an opportunistic rather than a purposive sample in that: although in some cases I could express a preference, I could not choose any of the vivas to be observed – they emerged as a result of who was willing to take part; and during the data collection period, the participating universities changed. Figure 4.2 shows a diagram of the cases, including the date of each observation and the institution where the viva took place.

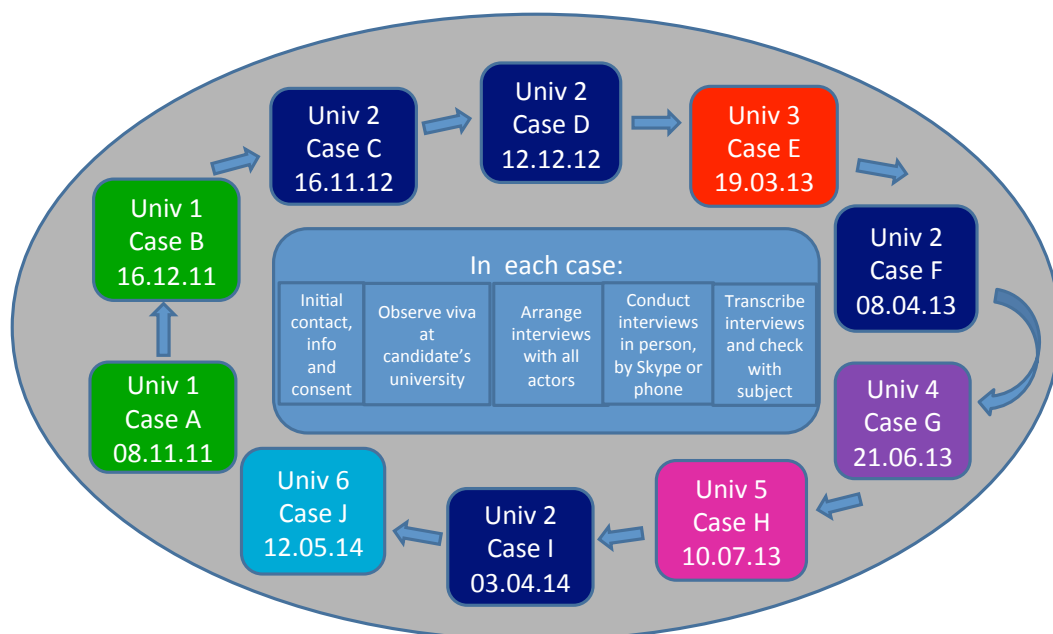


Figure 4.2
Ten bounded cases:
Observations and semi-structured interviews, November 2011 to May 2014

The cases are almost equally spread across science, technology, engineering and mathematics (STEM) subjects and arts, humanities and social sciences (AHSS) subjects, two broad groups

that encompass the majority of disciplines studied in UK higher education. Although candidates' anonymity is protected, their subject area is mentioned to emphasise points being made by respondents, for example, where the field of study affects the view being expressed. The disciplinary fields are listed below to demonstrate the range of subjects included in the study:

- Biosciences (dental microbiology, genetics, immunology and primatology)
- Education
- Film studies
- Media policy
- Sociology

The subjects of the experienced examiners interviewed (unrelated to viva observations) were: computer science, earth sciences, economics, mechanical engineering and molecular biosciences.

Quotations in the findings chapters (5 to 7) use a coding system to identify respondents. For case-related contributors, the prefix STEM or AHSS is used to denote the broad subject area, followed by a letter or letters to identify the role played in the process, as outlined in table 4.3.

Letter(s)	Role
E	External examiner
I	Internal examiner
C	Candidate
S	Supervisor
D	Main supervisor / director of studies
IC	Independent chair / convenor

Table 4.3: identification of roles

The numbers used in the codes broadly represent the order of the cases, but not always, because not all actors were interviewed and the examiners did not match the 'one external, one internal', model in all cases.

Non-case related contributors are identified first by their broad subject area (BIO, COMP, EARTH, ECON, ENG), followed by Ex (for experienced examiner) and either 1 or 2, depending on how many examiners in this subject were interviewed.

The final number of ten observations and associated interviews, combined with the expert interviews and focus group discussions, provided data that included 43 interview transcripts and ten case summaries.

12 Viva observations

Field notes from viva observations contributed rich data to the study and were the source for the case summaries in Appendix 1. Observations provided insight into both the process of the examination and the approach of the actors involved (Hartley and Fox, 2004). Case observation data on the whole supports and strengthens interview responses. References are made in the text of the findings chapters to indicate this.

12.1 Pilot

The practical challenges posed by the viva observations emerged during the first two cases at University 1 in November and December 2011. As pilot cases, these enabled testing of the observation protocol and interview schedules for examiners, supervisors and candidates and made a valuable contribution to the data while enabling refinement of the research questions and design. In the November 2011 viva it was possible to interview both examiners as well as the candidate and the supervisor very soon after the viva, in the case of the examiners the same day. One of the benefits of this was that the examiners had a clear recollection of how they had gone about judging the candidate's achievement and of the process that had just taken place. After the observation of the December viva, only the internal examiner and the candidate's supervisor were interviewed, some time after the viva. The pilots also confirmed the viability of the observation protocol and interview schedules.

The observation protocol

As agreed with the participating universities I remained silent during vivas. Some examiners, supervisors and independent chairs were most particular about where an observer should sit; others left it to me to decide. I had thought about this in advance and ensured that if possible I sat out of sight of the candidate. On the two occasions when I was visible to the candidate

(because of the size or configuration of the room), I sat at the side. Figure 4.3 shows the position of the individuals in each viva.

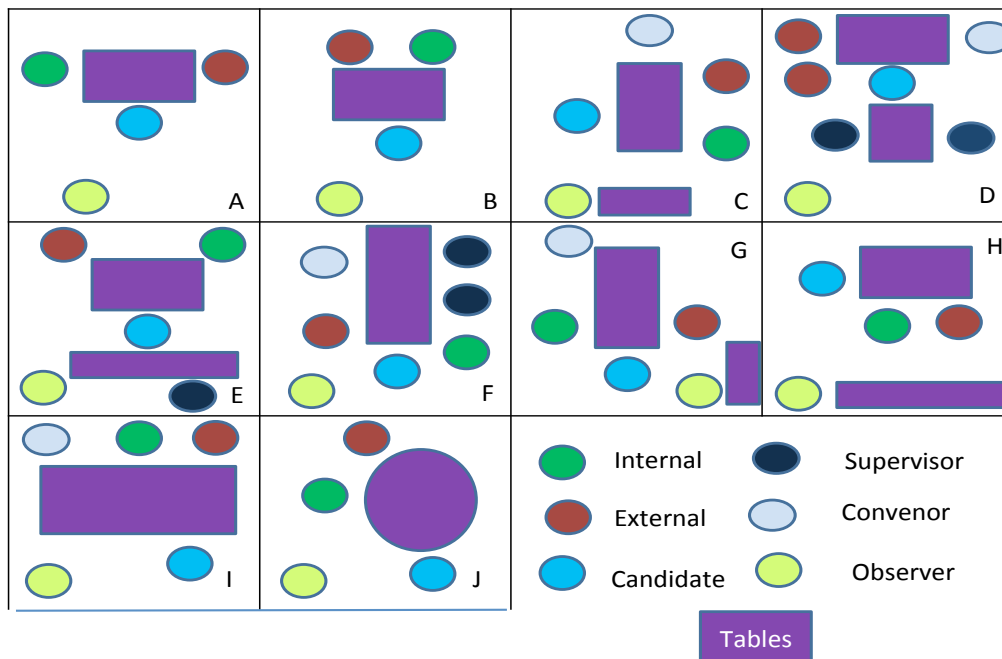


Figure 4.3: Positioning of individuals in each viva observation

It was essential, ethically and for the purpose of the study, not to distract either the candidate or the examiners. The viva needed to take place just as it would have done without an observer present. The conditions offered to all those agreeing to my observations included the assurance that I would not:

- participate in any way in the examination;
- record the proceedings; or
- make any noise that could distract the candidate or examiners.

I took care to fulfil this assurance and my attempt not to be in any way distracting was confirmed in an email from an independent chair to a colleague in his/her institution: ‘...Gill Clarke observed [a] viva I chaired and she was totally unobtrusive’.

In all cases I took detailed notes during the examination, even though in most cases the candidate’s topic was unrelated to my own knowledge or understanding. These notes have proved valuable in comparing the different vivas observed and are the source of the case summaries (Appendix 1). I also noted the length of time each viva lasted, for comparison between subjects and to consider if the viva length might be related to the extent and intensity of examiners’ questions on the thesis.

Apart from the likelihood that examiners and supervisors only agreed to my observing anticipated 'non-problem' vivas, a potential weakness in the design of this study is the possibility that the presence of an independent observer led to the actors behaving differently during the oral examination. For example, depending on their approach, examiners might treat the candidate more gently or harshly, or have evaluated the thesis more thoroughly than usual to show they had been diligent in their preparation, or the candidate might be more nervous because of the observer. It was not possible to evaluate whether behaviours were different, particularly with regard to the examiners' preparation. However, it appeared that during each observation, because of the intensity of what was occurring and the concentration on the examination by both the candidate and the examiners, that after the first few minutes of introductory remarks and especially once the candidate had been asked the first question, the others present forgot an observer was present. All were intent on the progress of the examination itself and, because I was silent, they all seemed unaware of my presence until the point where the candidate and I left the room for the examiners to consider their decision.

This is a critical part of the judgement where examiners discuss whether their respective views at the end of the process are similar and normally create a summary of the most important feedback on the thesis for the candidate. Examiners' discussions lasted for a different amount of time, partly because of the variable amount and complexity of the corrections they were asking the candidate to make but also because the procedure in some institutions included a requirement for the examiners effectively to write their post-viva report during this period. This approach seemed to enable examiners to be clear in their joint feedback to the candidate on the day when setting out what their recommendations would be; the formal written instructions for making corrections that the candidate would receive post-viva should not contain any surprises and should be received soon after the examination. Conversely, writing the report on the day in a rather pressured situation knowing the candidate was waiting to hear the result, could be argued not to give the examiners much time for reflection before formalising their joint report, rather than having the opportunity to iterate before sending it to the candidate.

13 Interviews

Semi-structured interview schedules (Appendix 4) were used systematically. If an interviewee said something of particular interest that I had not previously considered, supplementary

questions were asked to deepen understanding of the point. Most PhD examiners are also supervisors and the majority of supervisors have examining experience, both internally and externally. A single interview schedule was therefore developed that could be used with both groups, but which included questions about initial impressions and judgements that were different depending on whether the interviewee had been a supervisor or examiner in the relevant case.

All the face-to-face and Skype interviews were recorded, but not the telephone interviews. However, I took detailed notes during the telephone interviews, including some quotations. In this study, with its realist theoretical perspective, quotations were important in the data analysis chapters, for transparency and to demonstrate that individual views had been heard and were given parity with the researcher's interpretation. Table 4.4 shows the number of examiners, supervisors, candidates and independent chairs/convenors from whose interviews evidence was generated, together with the number of actors in each category present in the case-related vivas but not interviewed.

Case-related data						Non case-related data	
Inter-view-ed	External examiners	Internal examiners ¹	Supervisors ²	Candidates	Convenors/ Independent chairs ³	Experienced examiners	Focus group
Field							
STEM	5	5	5	4	2	5	1 ⁴
AHSS	3	4	4	5	1	2	
TOTL	8	8	8	9	3	7	Total inter-views: 43
Not inter-view ed	3 ⁵	0 ⁶	3	1	2	0	Total not inter-viewed 9
Notes: ¹ One internal examiner was also a supervisor in another case and was interviewed only once regarding both cases. ² One supervisor was a main supervisor (director of studies) for one candidate and a second supervisor in another case and was interviewed only once regarding both cases. ³ Convenors, or independent chairs, of vivas were used in five of the ten cases. ⁴ The focus group was made up of five supervisors / examiners in both AHSS and STEM subjects. ⁵ One of the external examiners not interviewed was the second external examiner in one case. ⁶ In one case there was no internal examiner; two external examiners were present.							

Table 4.4: Summary of interviews

14 Data Analysis

The complexity of the data required experimentation with different approaches to coding when deciding how the findings would be presented. Using the case-related PhD examination primarily as the unit of analysis (Bryman, 2008), I developed a coding framework (Appendix 5) using an approach that reflected the semi-structured nature of the interviews. The questions in the interview schedules were grouped under three broad headings: criteria used to make judgements; outcomes of the assessment; and weighting of the thesis and viva, but emerging data did not fit neatly into these categorisations which had not included consideration of the examination process as a concept. Rather than being constrained by the structure of the interview schedules, further development of the coding framework was driven by themes emerging from the data, namely: the assessment process; evidence used by examiners to evaluate candidates' achievements; and the final outcomes of the examination. These three areas best framed the analysis of the evidence provided by the richness of the research data and were therefore chosen as the main focus of the findings chapters.

14.1 Coding framework

Interview transcript and observation data were grouped under the three main chapter headings using long and short codes (Miles and Huberman, 2004), with sub-groups at up to two further levels to reflect fine distinctions made by respondents. For example, in coding data provided by examiners on the role of the thesis (highest level), there would be multiple codes representing the stated roles, e.g. as primary evidence for the judgement (middle level), then further codes (lowest level) representing how examiners use the thesis. This example is illustrated below as extracted from the coding hierarchy in the framework:

```
Role-thesis
  Thesis primary evidence for judgement
    Has-candidate-done-enough-for-PhD
    Does-thesis-meet-examiners-core criteria
```

Development of the framework supported data analysis and structuring the findings chapters. Coding was refined in parallel with the data analysis for each chapter. I began with data relating to the examination process, which included observation notes. This proved complex because of the close relationship between the examination process (Chapter 5) and the attributes sought by examiners (Chapter 6). Responses had tended to conflate the two areas so data had to be carefully analysed and separated. This in turn provided greater focus on the topic of the first findings chapter concerning the examination process. Developing the

framework iteratively highlighted the complexity of the data while also contributing to the analysis. It raised and resolved challenges encountered in assigning evidence to sections and chapters and deciding what data was significant and should be included or omitted, while ensuring the findings chapters were tightly focused on relevant data.

14.2 Data generated from interviews and observations

Interview data forms the most significant contribution to this study, supported by researcher observations. Table 4.4 shows the number and distribution of interviews conducted. Case-related interviews were dominated by examiners (17), but also included nine supervisor and nine candidate interviews. Convenor interviews focused more narrowly on the role of the convenor or independent chair in the process. A wide range of examiner views was represented, and on some topics such as the role of candidate publications in the thesis, clear differences emerged, depending on the subject. On other important points, such as whether or not the PhD examination should be graded, there was considerable agreement among examiners, supervisors and candidates and also among examiners unrelated to the ten cases. In addition to interview and observation data, I sourced some institutional regulations to illustrate, for example, what guidance universities provide on the doctoral examination process and its outcomes, including how they define the attributes sought by examiners (Appendix 6).

Predictably, the research generated more data than it was possible to include, for example, supervisors' criteria for advising candidates on readiness of the thesis for submission, detailed differences between institutional research degree assessment regulations, and the extent to which examiners' final reports reflect their judgements as observed by the researcher.

15 Conclusion

The methodology and research design for this study were chosen with the aim of discovering how the PhD examination enables examiners to make judgements about candidates' achievements by investigating the assessment process, by determining the attributes sought by examiners and by exploring a range of examination outcomes. The findings suggest that the research questions were effective and that the evidence generated adds to our understanding while supporting existing theories. Some particularly interesting findings have emerged on the examination process and the relationship between the thesis and the viva; the evidence contributing to examiners' judgements; and opinions about whether the examination should be graded. Evidence covers several fields in numerous institutions, giving a wider perspective

to the findings, which nevertheless cannot be generalised. Multiple case studies produced the rich data evident in the findings chapters and in the case summaries. The combination of interviews and researcher observations helped to increase objectivity, allowing comparison of two perspectives, one independent of the examination. Generating evidence about the role of the thesis and purposes of the viva in examiners' judgements through detailed scrutiny of ten bounded cases, also focusing on the relationship between them in the examination process, has proved an innovative approach.

Chapter 5: The PhD examination process

Assessment at the PhD level may not (any more than it is in other contexts) be a rigorously scientific process, but it is not a lottery either. (Becher et al, 1994:137)

1 Introduction

The two-part PhD examination process in the UK is variously considered to be '*an enigma*', '*a secret ritual*', '*a process of mystification*' or '*a lottery*' (Murray, 2011:2; Denicolo et al., 2000:64; Baldacchino, 1995:71-72). The secrecy and uncertainty is intensified by the private nature of the assessment, particularly the viva. The process begins with an independent evaluation of the thesis, normally by one internal and one external examiner. Both are required to declare their independence. The examiners share their preliminary views on the quality of the thesis before assessing the candidate in the viva. Each PhD examination is unique, influenced by the individual nature and quality of the thesis, the field of research, examiners' approaches, candidates' output and viva performance, and variations in institutional guidance. The relationship between the thesis and viva is complicated and poses some interesting questions.

In this chapter I examine the process holistically, exploring the role of the thesis, the purposes of the viva and the relationship between them to address the following subsidiary research question: *In the examination process, what role is played by the thesis, what are the purposes of the viva, and what is the relationship between them?* The chapter contains evidence from 40 (93%) of my 43 interviewees and five focus group members, as follows: eight external and nine internal examiners; all nine candidates; four supervisors, four non case-related examiners, three independent chairs/convenors and three focus group members. I also referred to examination regulations from four universities.

The PhD examination is a '*high stakes*', summative assessment (Biggs and Tang, 2007:142; Lovitts, 2007:20; Bourke and Holbrook, 2013:407) that because of the process and its essentially idiographic nature, confronts conventional expectations of validity and reliability in assessment in higher education (Field, 2009). A particular challenge is to demonstrate consistency of standards and process, given the unique nature of each examination. Some universities have introduced convenors of vivas to address these issues. As illustrated by Chapter 3, the majority of research undertaken rarely explores the relationship between the thesis and viva during the examination process.

The chapter is organised around six topics. In section 2, I suggest the idea of a 'continuum of judgement' to describe the examination process. Section 3 explores the role of the thesis, using data generated by interviews with examiners and supervisors and in section 4 I examine the purposes of the viva from the perspective of examiners and candidates, while section 5 addresses the relative importance of the thesis and viva in the examination process and their interdependent relationship.

2 The examination process: a continuum of judgement

The stages of the assessment process – independent thesis evaluation, followed by shared judgements and examination of the candidate in the viva – suggest an interdependence of the two elements of the process (thesis and viva) that I have described as a continuum of judgement resulting in the final outcome. A continuum normally implies something that is uninterrupted in time or sequence, yet in this study it represents both time and maturation of judgement. The time period includes intervals during which the process of judgement may be interrupted or suspended.

The interval between the examiner receiving the thesis and the staging of the viva is likely to span a period of weeks, if not months (Parsloe, 1993). This allows individual examiners to decide when and how many times they read the thesis and leads to multiple approaches. For example, an examiner may read a thesis soon after receiving it, make some preliminary judgements, then pursue other work until further scrutiny is required nearer the time of the viva. Others may not read a thesis straight away and might postpone any kind of judgement until the date of the viva has been set. In other cases, and perhaps influenced by the nature of the topic or the examiner's confidence in the thesis, multiple readings may take place as time elapses between receipt of the thesis and the day of the viva. Mullins and Kiley (2002) uncovered multiple patterns in examiners' approaches to reading the thesis, some returning to it several times before making detailed notes. Crucially, all had read the thesis more than once. While there may be as many sequences as there are examiners (Carter, 2008), the data emerging from this study support the idea of a continuum which begins with an examiner's initial reading of the thesis and ends during the examiners' private discussions prior to disclosing their recommendations to the candidate.

During the process examiners may also be influenced by factors such as new publications on the candidate's topic or information regarding the candidate's supervisor or department. Tinkler and Jackson's (2004) explanation of a three-stage examination process involves stage 1 being evaluation of the thesis, stage 2 the viva and stage 3 the examiners' recommendation and its implications, whereas according to Murray (2009:16), the viva *'is the final assessment of the research, the thesis and the student'*. This study, which was not designed to explore post-viva events, suggests that an examiner's judgement develops over time and that ideally even if not universally there may be three broad stages in the continuum that occurs before the examiners' recommendation, as outlined in Figure 5.1:

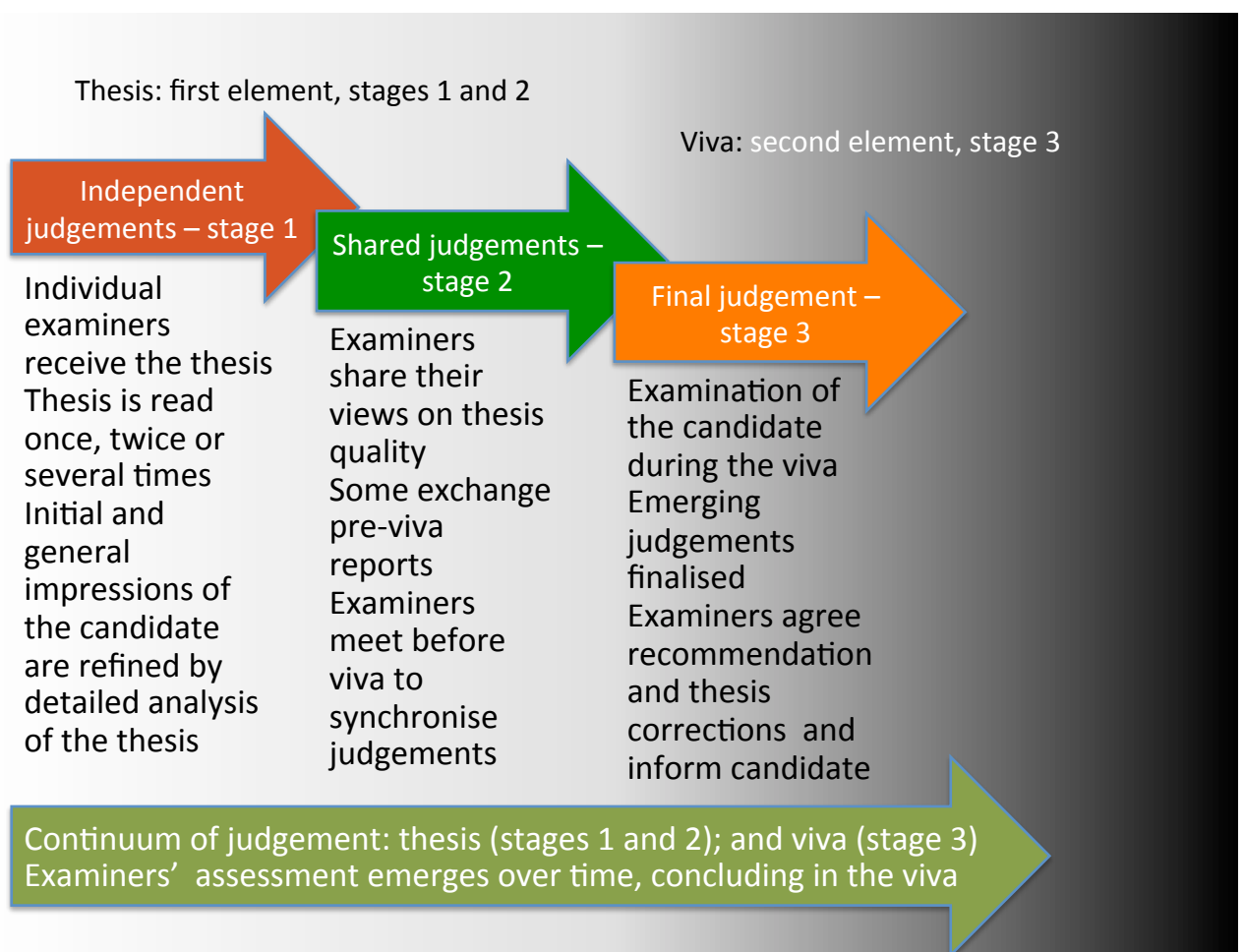


Figure 5.1: Continuum of judgement

The continuum begins with the judgement arising from initial evaluation of the thesis in stage 1, which may be refined as and when more in-depth analysis is undertaken before individual judgements are shared with the other examiner. In stage 2 individual judgements of the thesis may be moderated by the assessment of the co-examiner, either through the exchange of

individual reports in advance or during a pre-viva discussion on the day of the viva, or both. Most examiners interviewed for this study confirmed they usually broadly agreed with their co-examiner(s) about the thesis and the candidate's overall achievement. Stage 2 is the critical point where judgements are compared and synchronised: depending on the outcome of the pre-viva discussion an examiner may adjust their individual view or convince their co-examiner(s) to adjust theirs. The data in this study suggest that such an adjustment often involves relatively minor shifts of judgement, such as an examiner acknowledging that their co-examiner has noticed something they had not, or an experienced examiner paired with a novice examiner co-ordinating their judgements, each taking something from the other. During stage 2 judgements are synchronised as examiners become more focused in their analysis of the thesis while taking account of one another's views.

In stage 3, with the crucial dimension of input from the candidate, the examiners consolidate their judgements individually and jointly, the viva adding another layer to examiners' knowledge of the candidate and what has influenced the final outcome, including any difficulties encountered. Murray's description of the viva is helpful in this context:

The viva can be seen not only as a defining moment in the doctoral experience but as, in many ways, *the* defining element of doctorate research. It is the final assessment of the research, the thesis and the student. (Murray, 2009:16)

As suggested in Figure 5.1, the three stages of judgement are interdependent, with each stage influencing the next and, irrespective of the length of time involved, varying according to individual circumstances. Together the distinct yet interrelated stages provide a continuum of judgement, which begins with initial impressions arising from the thesis and ends in a joint, considered judgement after the viva.

Examiners' questions in the viva often appeared to me, the observer, complementary, suggesting that the process and in particular their discussion beforehand encourages a more reflective and balanced examination outcome than might be the case if only one examiner were involved, or if the process did not require examiners to negotiate their judgements. In most of the cases in this study, examiners had differing backgrounds, experience and perspectives on the topic, which they brought to the examination. In cases where there are no post-viva complications such as an appeal by the candidate, the last step in the continuum occurs when the candidate leaves the room once both examiners are satisfied that they have learned what they needed to know from and about the candidate; their discussion marks the

point of final judgement about the achievements displayed in the thesis when they agree the corrections to the thesis that are normally a condition of the recommendation to award a PhD.

Viewing the process as a continuum allows for both individual and collective judgements to mature and time for reflection. The second element, the viva, also gives candidates the potentially critical opportunity to defend their thesis. The contributions of the thesis and the viva to the continuum are explored further below, together with their relative importance in the examination process. Appendix 6 includes extracts from several universities' research degree regulations. Section 1 of the appendix concerns guidance on the process, mostly conduct of the viva. Some, for example, the University of London and the University of the West of England, make it clear that judgement of the candidate rests on both the thesis quality and the performance of the candidate in the viva. By integrating both elements, the London regulations make clear the dual and integrated contribution of the thesis and viva in the examination, as well as broadening the scope of the viva. However, regulations may emphasise the centrality of the thesis in the overall judgement, explored in the next section.

3 The role of the thesis

The thesis is central to the examination process. This is confirmed by institutional guidance on doctoral assessment (Appendix 6) and evidence from examiner interviews, which demonstrates the fundamental and practical role of the thesis in forming examiners' impressions of the candidate and their work (Mullins and Kiley, 2002). The thesis is also used by some examiners to benchmark candidates with a range of other individuals and their work in the field, similar to a peer review process. This part of its role is explored here in the context of addressing consistency of standards. I conclude by describing how the thesis acts as a reference point for the viva, raising questions to be explored in the face to face examination of the candidate.

The thesis has a fundamental role in the PhD examination. It displays the candidate's research design, conceptualisation, implementation and findings and forms the basis on which they are or are not admitted to the 'academy' (Delamont et al., 2000; Lovitts, 2007). Institutional regulations for doctoral degrees may emphasise the significance of the thesis in the examination process by describing its role in demonstrating the candidate's research achievements, most importantly their claim to originality but also their ability to conduct independent research and to situate their study in the wider field.

3.1 The centrality of the thesis

The majority of examiners, supervisors and candidates interviewed for this study were in agreement about the thesis being the main reference point for examiners' judgements. Interview data confirmed that evaluation of the thesis is the critical element of the process where judgements crystallise, even though they are fine-tuned in the viva, confirming the widely held view that the thesis is pre-eminent in the PhD examination. As this external examiner confirmed, the thesis demonstrates many of the attributes that indicate if the candidate has reached PhD standard:

I look and see whether the candidate will produce good figure legends, for example, and the quality of...tables...has there been a breadth to...the thesis, or is this a project which has been...focused...around one technique and the student has become a real expert in that technique and [applied it] to many...different samples, but really hasn't gained an awful lot of experience...of...experimental procedures?...Those are the...things that I would consider...[as well as] the quality of the literature review, the quality of the referencing, have the methods been presented in an appropriately scientific manner...could someone pick up that thesis and use the...methods chapter to...repeat the work? (STEM-E5)

This examiner deduces a wide range of information from the thesis, seeking evidence of: rigour in the candidate's research methods and depth of experience of scientific techniques; well-presented knowledge; and whether the project is appropriately situated in the wider context of the literature review showing a broader knowledge and understanding of the field. The examiner added that the thesis presentation should not occupy time in the viva so that discussion with the candidate could focus on the research. The fundamental attributes mentioned suggest the examiner looks to the thesis to indicate if the candidate has reached a satisfactory standard to be awarded a PhD, suggesting its centrality to judgements.

This example, together with other examiners' responses, supports the contention that the thesis is the primary evidence in deciding if a candidate has reached an appropriate standard for the award of a PhD and that detailed scrutiny refines examiners' thinking about the quality and presentation of the candidate's research that is then tested in the viva.

3.2 Initial impressions

First impressions of the thesis inevitably influence examiners' initial judgements (Mullins and Kiley, 2002). Even if subsequently judgements became more nuanced, some examiners in this study suggested that these early opinions, especially of the early chapters such as the literature review, affected their whole approach to the thesis and the candidate:

Yes, I think...early on in the thesis...the whole way in which the work is tackled, the introduction, the grasp of the field, starts to get you...either very anxious or very confident about the outcome. It doesn't take long for the thesis to start to convince or not to convince. (AHSS-11)

This suggests the pivotal role played by the thesis early on and that general impressions are sometimes simultaneous with gaining a deeper understanding of the candidate's research achievement. Initial impressions are mainly derived from the thesis and give examiners an opportunity to assess whether a candidate broadly meets their expectations for the award of the PhD.

Examiners interviewed for this study were encouraged to analyse the process in some detail, especially their approach to assessing the thesis. Some suggested that their initial impressions provided a good indication of the final examination outcome; others were less certain. Most examiners' initial impressions were based mainly on high level criteria, for example, overall coherence and a lack of obvious errors, complementary to but not as specific as some of the detailed or 'core' criteria explored in Chapter 6. The majority made some kind of distinction between initial impressions based on the thesis and their detailed examination of the candidate, combining judgements from both thesis and viva and using criteria relevant to the field, suggesting that judgements are refined and deepened by increased insight as the process progresses, with the thesis raising questions to be explored in the viva.

I asked examiners if they took an overall view on first reading, giving as examples the categories 'exceptional', 'good enough to pass without major corrections' and 'borderline', or similar. Several were likely to decide early in the process whether the thesis could be assigned to one of the three general categories I had suggested, or something similar. Fifteen of the examiners interviewed said that they do have one of these three, or similar, categories in mind on initial reading of the thesis. One external examiner commented: *'Of course, one...has a sense that some theses are much better than others...I would have a sense of ones that I was more confident about and less confident about'* (AHSS-E1). Another external confirmed: *'Yes, that would be one's thoughts – whether they're exceptional, solid pass or borderline.'* (STEM-E4). Yet another highlighted the third broad category as most likely to have a practical bearing on the final assessment outcome:

If they're borderline or likely to be asked [to make] major corrections...that will be in my mind when I'm thinking about the final assessment...The other two [don't] make much difference in terms of the decision, but that final one, when reading [the thesis], I'm

making a judgement of how much work I think might be involved in the revision and also looking at the criteria. (STEM-E2)

Two internal examiners commented explicitly on borderline theses:

I think you quite quickly have a sense, starting to read a text...not necessarily about...[a thesis with] no corrections...a stellar pass, because that depends on the other examiner [taking a similar view], but...you quickly spot a borderline. (AHSS-I2)

When you complete the thesis, even sometimes with three-quarters of [it], you start having an idea this is a really good piece or it's just something that...it's okay but...will not change the world – it will pass. I very rarely think that a PhD thesis is borderline. (STEM-I5).

Examiners' early impressions of the thesis may give an indication of their final recommendations. Responses also suggest that a borderline thesis is a somewhat rare occurrence, but when it occurs is easily identifiable and requires detailed scrutiny to determine implications for the viva and overall outcome.

An experienced external examiner thought the first category of the three was more significant, linking exceptional performance with the 'core' criteria used to make a judgement:

It's really important to note whether the candidate is exceptional or not...that...becomes clear very early...the exceptionality could be a number of things [including]...the originality of the argument is such that it stands out and that, for a PhD, is very important...in practice cases, the text involved... is in itself exceptional because of its...cinematic qualities and it stands out...again [that] is very important...a number of people who submit work quite clearly are not...exceptional...You read and look at the work and then it's a difficult judgement. (AHSS-E6)

This quotation suggests the few exceptional candidates are easily recognisable, for example, because of qualities of originality. According to respondents, most candidates produce a thesis that is good enough to pass with minor corrections. This was confirmed by the only examiner who deliberately did not take a view on initial reading of the thesis:

No, I don't take a view about whether the candidate is exceptional. I don't make any judgement. I would start from "Is this good enough to pass without major corrections?" Most people I've examined fall into this category. (AHSS-E3)

Examiners' responses suggest well-developed and thought-through criteria for the structure and content of the thesis that they would expect candidates to fulfil, and that they apply these throughout the continuum of judgement. Some suggest that in addition to having in mind general expectations for structure and content, on the first or second reading they are applying detailed criteria to check that the foundations on which the thesis is based are sound

and that any hypothesis has been rigorously tested. For example, the internal examiner below described how, in addition to having the three broad categories in mind on initial reading of the thesis, detailed content was of critical importance:

I do make a judgement...[based on]...structure of the thesis...whether it seems to be...a coherent piece of work...[either] the first or the second time [of reading]...the quality of the experiment is probably...the single most important thing...whether the experiment described makes sense, whether they're controlled properly...and whether the question that is being asked is somehow sensible. (STEM-I3)

In stating that the most usual opinion they had after reading the thesis was that the candidate was likely to pass 'with minor corrections' or 'without major corrections', these examiners (and STEM-E2 above), indicated they already had in mind the end of the continuum. Their ability to judge the final outcome at this stage was supported by viva outcomes.

Emphasising the significance of the summative element of the examination, an internal examiner argued for not assuming the candidate would pass, irrespective of thesis quality: *'Of course you make a judgement...you have to. That's straightforward. There's never an assumption they have passed and there's always going to be an assessment of how good it is'* (STEM-I2).

For others, giving candidates the 'benefit of the doubt', was considered important, for example this comment by a non-case related examiner: *'You begin by reading the thesis, giving them the benefit of the doubt – if it looks well-written you assume it's going to be good'* (BIO-Ex2). A similar approach was described by a case related external examiner who, agreeing that the three broad categories were borne in mind during first reading of the thesis, emphasised the importance of balancing positive and negative impressions in the overall evaluation, and compared it with the review of an academic paper:

Whenever I mark anything...I...want to be positive so I...start with the assumption...this is going to be great and definitely good enough to pass...then, as I'm reading...,I... clock up things that are not great and...start to weigh how significant those things...are. But...they can be weighed out by things that are really good and...at the end I make a judgement about...the things that are wrong with it – how significant they are...it's...important to bear in mind what is reasonable to expect in a 3-year period so I'm more lenient on a PhD thesis than I would be on a published paper... (STEM-E4)

In describing their attempts to be fair to candidates, the two examiners above present a constructive approach. Examiner STEM-E4 introduces another dimension: comparison with other authors. This is unsurprising given that most PhD examiners are authors and reviewers

of academic papers and that the PhD examination originated as a test for entry to 'the academy'; some contend that it remains focused on this concept (e.g. Denicolo, 2003).

Examiners' responses suggest that, using their experience and subject knowledge, from the thesis they assess how far the candidate demonstrates the level of research attainment needed to meet the criteria to be awarded a PhD. In applying their expertise, examiners are consciously categorising the candidate according to what the thesis reveals. Implicit in these judgements is the concept of 'originality' or a 'contribution to knowledge', but initially examiners appear to be gaining a high level overview of the candidate's abilities by analysing the thesis as a whole, rather than focusing exclusively on the novelty of the findings.

Most doctoral examiners are operating in an environment of continuous peer review and, although assessing a thesis is different from reviewing a paper written for publication in a journal, there are some similarities in that an individual is presenting their work for consideration and approval by respected researchers in the field. Some examiners referred directly to this process, which for the purpose of this study is described as 'benchmarking'.

3.3 Benchmarking

Respondents suggested that benchmarking candidates occurs relatively frequently, whether the comparison is with themselves, peers or other candidates they have examined. When examiners were asked about this practice, it was apparent that benchmarking was taking place either explicitly or implicitly. Some examiners, even before they were asked, compared assessing a thesis to reviewing academic papers. For example STEM-E4 above and the external below illustrate this kind of benchmarking:

I do [use broad categories]...but...it's very much a...mental view on the quality of the candidate through reading the thesis...So it's...a...mental tick-list that...I've produced, also you could say, from reviewing scientific papers...as well as examining.' (STEM-E5)

Two more examiners, having confirmed they had the three broad categories (exceptional, good enough to pass without major corrections and borderline) in mind when reading the thesis for the first time, also cited comparison with others among the criteria used to make such a judgement: '*By comparison with other students I've examined or supervised; by comparison with colleagues and with published work I'm familiar with; to what extent the thesis is a piece of original research*' (STEM-I1); and '*When I have the thesis in front of me I*

would also think about how it ranks in comparison with others and how many corrections might be needed' (STEM-E1).

Two non case-related examiners agreed that on reading the thesis they tended to benchmark the candidate with others and to approach it in a similar way to peer-reviewing. The first suggested that:

At one level it is like reviewing papers for publications...It's a question of evaluating [the] quality of what's written in front of you and doing some analysis and you either say 'Wow, gosh, how exciting' or 'Does the argument hang together well?' and...it depends on how the thesis is pitched and what it's trying to do. (EARTH-Ex1)

The second, for whom this was the first external examining role, confirmed having in mind the three broad categories when reading a thesis, referencing their academic experience:

It's a field...I have had some involvement in...I benchmarked them against myself...went through the thesis with a fine tooth comb and picked out various points for further questioning...comparing it with my own work and...other work going on within the department and equivalent departments...We read the thesis, we compare it to our own experience, our own...candidates...We say yes, this is good enough or it's not and I want to know more about this aspect of it. (ENG-Ex1)

As well as comparing the candidate's work with the work of others, this examiner has articulated a process that is more often implicit in our understanding of the PhD examination: benchmarking the candidate's work with one's own. This raises the question of consistency among standards in PhD examiners (Lovitts, 2007).

Five examiners confirmed that their professional culture of reviewing academic writing is brought to bear on a candidate's thesis, suggesting that some form of benchmarking takes place at stage 1 in the continuum. Examiner interviews and viva observations indicate that if the thesis is high quality, discussion during the viva resembles peer-to-peer discussion, especially if the candidate is explaining some of the technical elements of the findings. Where this apparent equality of status occurs, it supports the professional benchmarking approach described by these examiners – comparing research output with the work of peers.

One external examiner took the benchmarking point further by suggesting that some overall benchmark was needed for the PhD, yet arguing that adopting practice from elsewhere could potentially diminish quality. They found it *'extremely worrying, both as a supervisor and as an examiner'* that *'we don't have objective benchmarking for PhDs'*. Nevertheless, the examiner did not suggest how this might be achieved and was critical of benchmarking used in other

systems, suggesting it was *'artificial'*, being based on *'set numbers of publications ...where people will just write large numbers of poor quality publications to get the benchmark'* (STEM-E3). This was not an area discussed other respondents but does relate to the question of whether the standards of PhDs should be differentiated.

Not all examiners suggested they benchmarked candidates with other candidates and colleagues, for example:

I try not to [benchmark the candidate] because everybody is...different and the PhD is completely different. There's no standard PhD [or]...thesis that you produce. So I try (obviously it's very difficult not to, in some way, compare across theses, across people)...[to] take it as an independent piece of work. (STEM-E2)

In explaining why they tried not to benchmark candidates, this examiner articulated the idiographic nature of the PhD and the difficulties of demonstrating validity and reliability in its assessment (e.g. Biggs and Tang, 2007 and Field, 2009), referencing the individuality of each candidate and thesis. Nevertheless, they acknowledged prior experience could make some element of comparison difficult to ignore. It could even be argued that the use of benchmarking with other candidates or peers is desirable in helping to achieve some consistency of standards in PhD assessment within the relevant field, without compromising fair assessment – the concern of STEM-E2.

Some respondents indicated that the participation of more than one examiner, in itself arguably a form of benchmarking, helped to strengthen confidence in the examination outcome. One candidate suggested the involvement of an external examiner was *'absolutely essential'* (STEM-C4) to the rigour of the process. An examiner included the candidate's supervisor in the equation: *'so you've...got three different people all looking at the thesis and judging it, and so I suppose...that gives you some confidence that it's not just [an individual's judgement]'* (STEM-I3), suggesting that at least three individuals could have similar views about thesis quality. Another suggested that confidence in the judgement *'comes from the agreement...between the two examiners'* (AHSS-I3), another form of benchmarking. One candidate's confidence in the examination outcome was strengthened by *'the experience of examiners in their field and the knowledge and skills of the student. That's why people have two or more examiners'* (STEM-C3). In addition to emphasising the importance of shared judgement, this candidate's comment suggests that examiners' experience in the field of research provides insight into the standards required, implying benchmarking through shared

understanding. These comments suggest that confidence in the examination process may be linked with the standing of examiners in the field and their research experience.

Data in this section indicate that examiners develop 'internal' benchmarks through varied experience that they use in the PhD examination. These may have emerged from: assessing other doctoral candidates; their own research; or evaluating the work of peers and other researchers. The thesis analysis, including any benchmarking that takes place, leads to a profile which cannot be deepened until the examiners meet the candidate. It is key to identifying questions to be explored with them in the viva. Several respondents gave examples of how viva questions arose from their evaluation of the thesis.

3.4 Raising questions for the viva

The quality of the thesis has a significant impact on examiners' preparations for and approach to the viva, as suggested by this non-case related examiner, who confirmed that the thesis affects the examiner's way of thinking in the approach to the viva: *'If badly written it makes the examiner more probing – if the thesis is sloppy it makes you think they might also be sloppy with the science'*. (BIO-Ex2).

Viva observations confirmed the role of the thesis in influencing viva questions and informing examiners' thinking about the nature of corrections, with many references to page numbers, figures and thesis chapters. It was also clear that some theses raised numerous questions about substantive matters, for example: methodology, including: sampling; statistical methods and data interpretation; and whether the conclusions were borne out by the data. In some of these cases, greater clarity and rigour in writing the thesis might have led to examiners approaching the viva with increased confidence about the candidate's achievements (as suggested by examiner AHSS-I1), but in others the questioning arose from, for example, the examiners' wish to understand fully a candidate's topic, experiments and interpretation in a complex and detailed study. This latter purpose often relates to the examiners' deep interest in the candidate's research and is unrelated to corrections (e.g. the examiners may have decided on the basis of the thesis that corrections would be minor). Yet a degree of uncertainty remains until the examiners and candidate have met and explored the thesis content in depth.

Examiners emphasised the continuum of judgement by describing the integrated nature of the process and confirming the role of the thesis in viva preparation, as demonstrated by this non-case related examiner:

From the thesis...with certain modifying factors based on the performance of the candidate in the [viva]...I would have a broad idea about the standard of the candidate before meeting them...based on the technical content or the presentation, or whatever aspect of the thesis. I... would target some of the questions...to try and establish whether I was correct or not in those assumptions. (ENG-Ex1)

An internal examiner, describing their method of evaluating the thesis, demonstrated a systematic approach to determining viva questions: *'starting at the beginning...I tend to make notes as I'm going through the thesis...which will highlight questions...I might want to ask'* (AHSS-I3). In describing how viva questions emerged from the thesis, examiners underlined its centrality to judging the candidate's knowledge and understanding, in some cases relating their comments directly to the case observed. In the following example where the candidate was also a practitioner, the structure and composition of the thesis was unusual, the candidate's research strengths shone through:

I was...surprised at the presentation of the thesis and although I could see quite clearly that there were some good data...I was concerned...there was no attempt...to place the results within the context of the relevant field...there was no discussion whatsoever and no general introduction...something I'd never seen before. So, I did go into the viva with a number of questions and I needed to understand the situation that the candidate was in during the write-up period and...the whole period of the project. (STEM-E5)

In this case the thesis raised some fundamental questions that were explored satisfactorily during the viva, leading to both a revised thesis and a positive examination outcome. The examiners recommended the candidate should pass with major corrections, to be completed within 6 months.

Two internal examiners provided examples of where literature references were missing from the thesis, for example: *'it's a topic...where I know the literature...well...there were severe gaps in knowledge...so you...formulate questions according to what you see is missing...'* (STEM-I5). Summarising several 'core' criteria on which judgements are based, the other internal noted the opportunity provided by the viva to pursue gaps in the literature: *'They may have missed some papers and you can discuss these with them during the viva'* (STEM-E1). These examples show both the importance of situating research in the field and the contribution of formative feedback from examiners, a great deal of which was identified in the cases observed.

The pivotal role played by the thesis is illustrated by data presented in this sub-section. Not only does it, in many cases, allow examiners to be confident that the candidate has '*done enough*' (ECON-Ex2) to merit a PhD, it forms the main reference point for the viva. From the thesis, examiners gain impressions about the candidate and their work that is often reinforced, rather than overturned, in the viva. The thesis also constitutes the evidence for examiners' negotiation of their joint preliminary judgements. It provides the reference point that enables the candidate's work to be compared with, or benchmarked against, other academic writing. Finally, and arguably most importantly, the thesis provides evidence of the candidate's originality or contribution to knowledge. Evaluation of the thesis evolves along the continuum of judgement leading to the examination outcome, informed by the candidate's defence in the final stage. Recognition of this central role, however, does not suggest the viva is rendered irrelevant. On the contrary, it enables the candidate to defend their work and helps to refine examiners' judgements.

4 The purposes of the viva

In the PhD, the term '*viva voce*' represents an oral examination of the candidate on their thesis by at least two examiners. This section explores the purposes of the viva through regulatory statements and candidate and examiner perspectives and observations, some of which contrast with views expressed in the literature. Appendix 6, section 1, contains examples of guidance on aspects of the viva. Some, such as the University of Leicester, provide guidance on the viva specifically for candidates. The examples in Appendix 6, while varied, are supported by Jackson and Tinkler's and Tinkler and Jackson's data (2001 and 2004 respectively) and Poole (2015), and corroborated by evidence from observations and interviews with examiners and candidates.

The complexity of the viva is illustrated by several characteristics. First, it is a time-limited and dynamic phenomenon that, unlike the thesis, gives examiners only a short period for reflection before finalising their judgement. It necessitates the integration of two (or more) examiners' opinions and it has to allow for the dialogue to take unexpected turns, depending on the candidate's responses to examiners' questions and the ensuing discussion and therefore subject to the uncertain nature of human interaction. The candidate's university regulates the viva procedure, which is therefore variable. Regulations may include, for example, the compulsory attendance of an independent chair or convenor as in five cases in this study, or rule that the candidate's supervisor may not be present even with the candidate's permission.

The examination process technically extends beyond the viva and takes account of different institutional approaches to the outcomes, such as the length of time taken to complete minor or major corrections. For this study, however, the viva constitutes the final stage of the examination.

The data in section 3 make a strong case for the pre-eminence of the thesis in the examination process, yet for some of the respondents the viva fulfilled several significant purposes. In summarising the rationale for the viva, one examiner succinctly suggested: *'It's to assess the suitability of the candidate for the award of the doctorate. To find out whether they really know their subject well and can give a good account of what their thesis is about'* (STEM-I2). Another summarised the main purposes as follows:

To ensure the candidate has undertaken the work [and] is conversant with the field; [to] provide an opportunity for the candidate to clarify aspects of the thesis, and to ensure quality control for the award of the degree...those are...key. (STEM-E5)

Others suggested more explicit purposes. Using recent experience as an external examiner, one supervisor argued that the viva was significant for shared examiner judgements:

I was asked to examine for an Australian university, which doesn't have a viva...I sent back my report, which asked for...major revisions, and invited the internal examiner to get in touch with me, but so far I've heard nothing...that couldn't arise here because you would meet at the viva. (AHSS-D2).

This supervisor underlined the importance of an examiner dialogue, implying the process has increased validity when examiners are able to communicate directly and agree a joint judgement.

Much of my data concerning the purposes of the viva aligns with others' findings. Purposes include: to provide the candidate with an opportunity to 'defend' their thesis; to enable examiners to confirm or modify their individual judgements and deepen their understanding of the candidate and their research; to act as affirmation for candidates at the end of their PhD programme; to provide feedback; and to confirm authorship.

4.1 Thesis defence

The concept of the viva as a mechanism for defending the thesis emerged from examiner and candidate perspectives. In its capacity as the final stage of the examination, the viva provides an opportunity for the candidate to respond to probing questions about their thesis, and for examiners to moderate their judgements, learning more about the candidate and the research

and coming to a shared conclusion. This and the next section explore the viva from both perspectives, beginning with the candidate. Several candidates used the word 'defend' in this context, for example:

It's probably to give you a way to defend yourself? In some cases there's a reason for doing something but maybe [you] didn't explain it too well in [the] thesis, although it might be reasonable, so [the viva gives you] a chance to explain it. (STEM-C2)

An AHSS candidate joked that one of the purposes of the viva was 'to scare you', then suggested more seriously:

It gives you a chance to defend the bits that are weaker...it gives you a chance to explain – when they challenge something you can say, 'Look, I've said something else here', or, 'But the reason it looks like that is because of this'...that...can be important when it's a marginal pass. (AHSS-C1)

By acknowledging that the thesis may contain weaker elements, these candidates make explicit a factor sometimes overlooked: that however good some parts of a thesis may be, the candidate may not have achieved similar levels of high quality across the whole thesis, or have provided clear explanations about the research results. These two candidates used 'defend' in the context of being able to explain or justify any weaker or unexplained elements. Two other candidates also used 'defend' in this context: the first suggested that without a viva there was potential for the examiners to '*Mis-assess you by judging things in your thesis which you don't have the opportunity to defend...Maybe if you explain it further in the viva...it makes sense*' (STEM-C5). The second candidate raised the added attribute of the ability to think quickly: '*Pulling in everything you know about the subject area and being able to defend it on the spot*' (STEM-C4). These four candidates suggested the viva fulfilled an important function in enabling them to explain their choices and findings. The next focuses on the opportunity to explore the research in more depth with the examiners, giving implicit emphasis to the concept of a defence:

To discuss and...test the candidate – if [their] theory or findings are really consistent....It's necessary for answering questions...from...reading [the thesis]...and for a proper evaluation...testing...the research...When you write the thesis you have constraints of time and space so there are lots of things that are not written there that the examiners could think you aren't aware of, or you didn't think of, or that [you] don't have a clue about...so...the viva is an opportunity to test [if the] researcher...knows the subject. (AHSS-C3)

This candidate approaches the viva from the point of view of 'testing' the candidate, who is required to defend what is written or omitted and how the research has been conducted. They suggest that the permitted length of the thesis and time constraints might inhibit the author

from explaining satisfactorily some details about the research and that the examination would be incomplete if the candidate and the examiners did not meet. Another candidate's response also suggested: *'There's always so much more that doesn't go into it, which even in a thesis...theoretically it's supposed to be everything but [there are a lot of things] you don't write up'* (STEM-C3).

These candidates implied their examiners would have had an incomplete understanding of their research study and outcomes were it not for the viva. Another candidate was aware beforehand that the external examiner *'Wanted to clear up'* some theoretical questions and also ask about some *'Omissions of things that probably should have been [in the thesis]'* (AHSS-C4). My observation, confirmed by the internal examiner, indicated this candidate provided a strong justification of the thesis, enabling them to defend the theoretical position taken and discuss with the external examiner why some information had been omitted. This example also implies that some prior knowledge of an examiner's concerns can help to prepare the candidate to defend the thesis. This is not normal practice in the UK, however, and is controversial. An informal interview I conducted with a dean of postgraduate research degrees in Canada indicated that by contrast candidates there are routinely informed of examiners' initial impressions and have an opportunity to address any concerns before final submission.

The evidence suggests that from the candidate's perspective, the opportunity to defend the thesis in person is a significant feature of the UK examination. Some examiners supported this view, for example:

It's important for there to be an opportunity for the candidate to be able to defend their work face to face...by word of mouth...and I'd be unhappy about failing somebody without giving them a chance to explain what they're doing. (STEM-I3)

Another reasoned: *'To understand...choices [the candidate] has made...to clarify any misunderstandings, you've got...to push them on their position and get them to defend those choices'*. (AHSS-I2) An external examiner asked:

Have they got an ability to...engage in [the] kind of dialogue that is characteristic of the ...viva where...people challenge what they've done and they've got to defend it...has this candidate understood the questions...are they making a reasonable defence? (AHSS-E1)

Another examiner suggested that in the UK the viva had become *'a sort of ritual'*, more like *'a scientific discussion and not so much an exam any more'* (STEM-I4), claiming it was no longer a real test of the candidate because so few people failed. This examiner added: *'My impression*

is it's not any more a defence because...we were always very nice with the candidate. We never...attacked the candidate for them to defend the work' (STEM-I4). This examiner was comparing the UK system with the continental public defence, also raising the question of exactly how challenging examiners should be in the UK viva without intimidating the candidate, a topic that continues to attract media attention.

As the literature suggests, the extent to which the viva is successful depends on several factors, including the candidate's preparedness, and the examiners' approach. The next response indicates that examiners' expectations of candidates, based on reading the thesis, are not always met:

It's a very...hard thing...a viva, so sometimes, probably, people can under-perform...you might be more impressed by the thesis and when you meet the candidate you can be quite surprised how difficult they find what...you might think [of] as quite gentle questioning...So yes, the viva does play a part. (AHSS-E1)

In implying that not all candidates are equally successful in defending the thesis, this examiner highlighted the need for candidates to be prepared to mount a strong defence in response to examiners' questions. Some candidates this examiner had encountered did not appear well prepared for the oral examination, finding questions about their work more challenging than expected. None of the candidates interviewed suggested their examiners' questioning had been intimidating or unreasonable, although some did admit to finding the experience demanding. As already mentioned, however, my opportunistic sample of ten cases was unlikely to include 'difficult' vivas.

4.2 Opportunity for examiners to confirm thesis judgements and deepen understanding

The candidate's defence enables examiners to test whether their thesis-based judgements are borne out and to finally judge if a candidate's achievements are worthy of a PhD, confirming or modifying judgements and deepening their understanding of the research. The latter was particularly evident in my observations, which revealed examiners' keen interest in candidates' findings.

Examiners are simultaneously evaluating the candidate's ability to defend their work and to explain any elements of their research not addressed in the thesis: *'For me it's about exploring why weaknesses have occurred and really trying to understand...is this just a failure of the student to explain this well or do they actually not understand this?'* (STEM-E4). It follows that examiners' impressions of the candidate from the thesis may or may not be confirmed during

the viva. In some cases, confirming prior judgements appears to have been a relatively straightforward process where the expectations of the candidate derived from reading the thesis were a good indication of the final outcome:

The thesis was longer than I expected when I received it. My preliminary report highlighted that the thesis was well-written and structured, making it easily readable and to navigate. It was of the quality I expected and the student bore out my expectations on the day. (STEM-E1)

Other examiners and supervisors confirmed their pre-viva expectations had been correct: *'Yes, my pre-viva impressions were borne out in this case'* (STEM-I3); *'Yes, my impressions were borne out in the viva'* (STEM-I1); *'In the recent viva, my pre-viva impressions were borne out pretty well by the end'* (AHSS-E3); *'Most of the time things turn out as you think they will. The one or two students I've had who I thought were very good came out of the viva with no corrections at all'* (AHSS-D1); and *'My answer would be that, on the whole, vivas do bear out previous conclusions'* (AHSS-D2).

The comments above might suggest that examiners would be prepared to make a judgement solely on the basis of the thesis, yet none of them considered that the correlation between pre- and post-viva judgements meant the viva was unnecessary and three other respondents alluded to its usefulness in confirming or modifying judgements, as follows:

The primary basis of the judgement is on the written material...that's how you form an opinion and then you modulate it...in the viva, because there are some things that you don't know from the written work about the candidate. (STEM-I2)

If [the thesis] looks well written you assume it's going to be good, then in the viva you have the opportunity to test whether the student is as good as [s/he] looks from the thesis. (BIO-Ex2).

The viva is an important part of the process...most of the examiners I've worked with...see...that it can tip the balance one way or the other. I don't think I've ever been in a situation where we've had a very...strong thesis and...a terrible viva...[but the situation] that I have been in is [one] where [the thesis is] on the borderline and actually they retrieve it and...argue it well. (AHSS-S2)

Case J in this study exemplifies this situation. While unequivocally assigning primacy to the thesis in making a judgement, examiners and supervisors acknowledged that the viva corroborates (or not) the thesis-based view of the candidate. Examiners implied they might be reluctant to make a judgement about a candidate without the benefit of the viva, confirming that they rely on the face-to-face meeting to: learn more about the candidate and their research; to understand their reasons for taking certain decisions; to discuss any technical difficulties; and to affirm or modify the initial judgement.

This candidate's view was that by demonstrating a broader and deeper understanding of their topic than is evident from the thesis, candidates can positively influence examiners' final judgements:

That you understand the subject area...and that you are aware of problems...but also ...strengths...being aware of problems and...to think of other possible ways of handling those...in the future and...problems that you weren't aware of, being able to...think of other possible solutions. (STEM-C4)

They also emphasised the need for consolidating knowledge and the ability to problem solve at short notice. Another candidate suggested: *'Probably [the viva is] to see that you understand and know what you've done, why...and how you did it and how you're going to take your research forward'* (STEM-C3). The same candidate commented further on how the viva can influence and confirm examiners' judgements:

They understood my work and my process better [after the viva] and why for example I didn't finish or go on with the first question that I had...why I focused on other things...You can always get so much more out of it by talking than if you just read the paper. (STEM-C3)

The viva enables the examiners to test the candidate's understanding of their own research as well as making a substantive contribution to the examiners' understanding of it, including explaining any difficulties encountered. Space in the thesis is limited and as noted by the candidate above, it is not always possible to include details about unexpected complications. Another candidate provided the following summary of the ways in which the viva can provide examiners with additional knowledge about the candidate, their strengths and weaknesses:

The purpose of the viva is...to merit a PhD it's not just about the quality of your writing, it's the quality of your presentation and...your embodiment of knowledge or...research and how to express...or communicate that...It's possible to write things in such a way that you're putting your best foot forward but masking...gaps in knowledge...The interaction possible in a viva is very important, to test...the learning and dig underneath the surface of the [thesis] to see if you truly understand or appreciate the questions...underpinning the work...It's...an important validation of what seems to be...in the written [work]. (AHSS-C5)

These candidates acknowledged that discussing the thesis face-to-face gives the examiners a better understanding of the research as well as of the candidate's wider comprehension of their research results. The last candidate also pointed out that it is difficult to cover up any lack of knowledge in the viva. From candidates' perspectives the viva is an opportunity to mention elements of the research it was not possible to include in the thesis, as well as to display

personal and professional skills such as the ability to communicate well, problem-solve and think on their feet. These viva characteristics were observed in several cases (Appendix 1).

Some examiners use the viva to explore the candidate's motivation and entry route. For example, a non-case related examiner referred to a recent viva in which they had been the external and where the overall judgement on reading the thesis was that: *'the written work was okay but lacked some depth'* adding *'I got the impression that [the candidate] was probably not going to be a strong student'* (BIO-Ex1). The examiner continued:

I started to dig around...[to] find out...about the person...I wanted to know how did they come to do this, why did they choose this PhD, why this supervisor? And it turned out [the candidate] had been working....[They] got a bit bored with being a [lab] technician and wanted to do something that stretched them...[so] decided to do a PhD'. (BIO-Ex1)

In this case awareness of the candidate's background provided additional insight into the quality of the thesis, which, although well written, did not demonstrate an in-depth ability to critique academic papers and suggested some problems with the 'nitty-gritty' of the research. Further questioning elicited the information that each year of the programme had raised challenges, leading to the candidate beginning another project after a three-month delay. In this (unobserved) case, the examiners concluded that a more productive outcome to the project might have been achieved if the candidate's supervisors had advised them to undertake a lower risk project. The viva provided the examiners with information they could not have deduced from the thesis and gave them a broader base for making a judgement, while corroborating their impressions from the thesis.

Another examiner described the need to understand the supervisor's input as part of the confirming or modifying thesis-based judgements:

There were concerns...about the structure of the thesis, the lack of consideration of how the results fitted with those in the field...the viva did clarify some of those issues, and there was quite a lot of discussion...to be able to understand the situation...the student was in...[I] always...ask PhD candidates whether they have any comments...about the supervision..., had they had access to the equipment...they needed, were there any concerns about access to consumables...the nuts and bolts of running a project...The discussion...with this particular candidate was helpful in that respect. (STEM-E5)

The previous two examples reflect the role of the viva in raising examiners' awareness and increasing their understanding of positive and negative influences on the candidate's experience and on their thesis.

The additional information gained through the viva provides a rounded perspective of the candidate and their work. Some argued this facilitates a more nuanced judgement and potentially a better-informed examination outcome. My respondents' comments align with Tinkler and Jackson's findings (2004) in that they do not suggest the viva affects the threshold judgement, but may have a bearing on the nature and extent of corrections to the thesis.

In commenting on the value of the viva in confirming or modifying examiner judgements, several candidates and examiners also identified it as an affirmation of achievement and 'rite of passage' (Denicolo, 2003; Morley, 2004; Tinkler and Jackson, 2004).

4.3 Affirmation of the candidate's academic achievements: a 'rite of passage'

For some, the concept of the viva as an affirming experience for candidates may not be its most obvious purpose. However, candidates' insight into the examination included recognising the positive impact of the face-to-face meeting with examiners, some describing it as a 'rite of passage', for example:

I...see [the viva] as a rite of passage and these people are...inviting you in and then you're finally an academic...It...feels good to have this...rite of passage and something that is actually quite scary to finish it off...It would be very strange to work on something for 4.5 years and then you just hand it in and that's it...It feels kind of good to be scared and then come out and be really pleased and...you...need to go through something like that to mark that you have done it and it's over now. (STEM-C1)

This candidate is linking the 'rite of passage' concept to acceptance in the academy and, in contrast with Poole's *'interviewee T'* (2015:96), sees this as a positive experience rather than something disagreeable to be undergone. In response to a follow-up question asking if the viva was more scary because the outcome was unknown, they continued:

Yes, because...even though...I knew that I'd pass...there was still a possibility that I wouldn't... Having done it, I feel better about it, so having gone through something difficult and...passed it...makes me feel better about the thesis, to know that...I actually knew a bit about my research. (STEM-C1)

Even this candidate, who defended the thesis confidently and appeared comfortable being challenged on elements of the research, found the prospect of the viva 'scary'. However, when the interview took place it was viewed as a positive experience that corroborated the achievement demonstrated in the thesis.

Interestingly, in this study, candidates' perception of the viva as a rite of passage contrasts directly with those interviewed for Jackson and Tinkler's (2001) research, who did not mention

the phrase. Another candidate in this study used 'rite of passage' to describe the viva, also commenting on the concept of entry into 'the academy' signified by the PhD:

This idea of it being an indoctrination, a rite of passage, because...the PhD is awarded only on the basis of that work or not, you can't fail somebody because you look at them and after they've defended their work say that they have no capacity to go on and be that productive member of the academy. (AHSS-C1)

Another candidate emphasised the feeling of affirmation, suggesting that the viva:

Is maybe not so much about the assessment of you as a PhD candidate but how helpful it can be to...be able to talk about your work and listen to what other people think about it. ...For that purpose, rather than just...assessment...it could be incredibly useful and...it was and wasn't for me at the same time, but should be. [AHSS-C4]

Both the AHSS candidates recognised that passing the PhD examination, as signified by successful completion of the viva, involved a process of affirmation and recognition of the work by their peers, a purpose related to benchmarking. The second also demonstrated that feedback may not always be entirely useful.

In describing the difference between an anonymous review of a journal article and the PhD examination, another candidate expressed appreciation of the affirming opportunity provided by the viva to talk with the examiners:

In the viva we...talk to the examiners face-to-face. It's quite different from when...you submit [a paper] to a journal and...reviewers review your manuscript and you feel very different...The viva is very helpful in itself. (STEM-C2)

This perspective was supported by the view of an external examiner who, without mentioning the words 'rite of passage', nevertheless implied the concept:

[The candidate has] done something amazing and spent years slogging their guts out, just to get to that point. It would be a shame just to get a print-out from the registry saying you've met the criteria. It's a celebration of their work. (STEM-E1)

In a wide-ranging discussion of the purposes of the viva, members of the focus group contributing to this study mentioned the opportunity to evaluate the candidate's 'ownership' (FG-F2) of their research, or 'looking to see whether they're an academic' (FG-F1). One contributor suggested 'There is, I think, an important ritual dimension to the viva' (FG-M1). Tinkler and Jackson (2004:27) also refer to the concept of the viva as a ritual: 'Whilst ritual is not necessarily the primary purpose in the UK, it can be a purpose in some cases'. They also suggest that the PhD examination and the viva in particular 'is feared and often poorly understood by many students' (p.2), but the nine candidates interviewed for this study showed

insight into its purposes and were largely uncritical of their experience, appearing to view the viva largely as some kind of affirmation rather than an ordeal. This cannot be taken as indicative for all candidates, however, because of the potential bias inherent in my sample, some of whom were also appreciative of the feedback provided by examiners, as explored in the next section.

4.4 Providing the candidate with feedback

The concept of the viva as a 'defence' implies a completely summative examination, such as in some continental systems where the thesis does not change following the assessment. This study, however, suggests that in the UK it combines formative and summative assessment and that feedback often leads to changes to the final thesis.

In the cases observed, it appeared that a significant amount of feedback was given during the viva, occasionally resulting in a requirement for the candidate to re-write some sections. While feedback is principally for the purpose of improving the thesis and linked to the nature and extent of corrections required, it may also extend to advice about professional skills, particularly for candidates known to be entering post-doctoral academic roles. Feedback I observed included: suggesting additional references to literature the candidate could or should have mentioned in the thesis; moderating statistical techniques and presentation of statistics; advice about publications - emanating from the finished thesis, yet to be written, or already included in the thesis; and advice about professional academic skills (such as how to present multiple authorship if a previously published paper was included in the thesis). Examples of the feedback provided to candidates are included in Appendix 1 and support data in this section, which explores the nature, amount and purpose of feedback provided to candidates.

In all cases I observed that examiners offered candidates detailed, sometimes extensive, feedback, most of which took the form of constructive criticism of the thesis and led to minor or major corrections to be completed as a condition of being awarded the PhD. Some examiners had marked up the thesis carefully with suggestions for corrections that were passed to the candidate at the end of the viva as an optional aide to revising the thesis over and above what was formally required and which would be summarised in the examiners' joint report after the examination. The nature of feedback may in some instances have been influenced by the examiners' knowledge of the candidate's future plans, for example if they had secured a post-doctoral position or were returning or progressing to another career. In

cases where the candidate was already in an academic role or planning to enter a post-doctoral research position, emphasis appeared to be given to academic matters such as: improving statistical knowledge and techniques; re-writing elements of the thesis to achieve greater suitability for publication, or to improve structure and coherence in the expectation that it would be read by peers in the field. Feedback in the cases where the candidate's career path was unknown, particularly if the thesis was of lower quality, appeared more focused on corrections that would improve the thesis sufficiently to meet the PhD 'threshold' while also enhancing its academic status. One external examiner in a STEM subject summarised the position of several examiners as follows:

The extent of minor corrections...might not be so stringent if the student isn't going into a science career and their research isn't ground-breaking. In that case...I wouldn't ask them to hone it or do a large amount of corrections. If going into a research [or] academic career a PhD student has to get their papers out as soon as possible after the viva. It is crazy to ask a student to spend weeks perfecting a thesis that probably no-one's going to read when they could be getting that knowledge out to a wider audience...examiners are likely to mention things in the viva that need to be put right before publishing. (STEM-E1)

This examiner seems to be suggesting that in either situation, they would not want a candidate to spend undue amounts of time correcting their thesis, but that it should simply be 'good enough' for the award of a PhD. Another STEM examiner acknowledged that *'the viva is useful for the student as a means of getting feedback, giving them ideas about how to go forward'* (STEM-E1), suggesting that one of the formal responsibilities of examiners is to provide the candidate with constructive criticism that may include formative feedback.

Most candidates appeared appreciative of the examiners' constructive comments on their work, perhaps suggesting that a characteristic of some doctoral researchers is the wish to produce as high a standard of research as possible. Some candidates were more inclined to voice their appreciation than others. For example, a STEM candidate's comments on the purpose of the viva included strong appreciation of the feedback received and the importance of being able to accept constructive criticism as a professional:

Colleagues and peers...can all help make your work better...than...if you just worked on it [alone]...it's training...If you go into academia and research...it's all about [passing] your work through peers, you have to deal with feedback...when you're getting anything peer reviewed...you have to take it in a constructive way because generally that's the aim of it...there's some accepting that that is part of the process and [the viva] is getting a taster of it....some outside party coming in with a completely different approach...a fresh perspective [and offering criticism], you have to...not take it defensively, just confidently defend your work and take their suggestions on board. (STEM-C4)

This candidate's perspective demonstrates the personal characteristics of maturity and professional experience, supporting evidence from other cases observed. Some candidates thanked the examiners during the viva for their feedback and advice on how to improve the thesis and/or on the skills that would be demanded of them in their career. For example, this candidate, when discussing the examiners' attention to detail, agreed that *'they concentrated on details, like if you want to be a scientist you have to be careful, even like citations or how many zeros you put in your data...it [is] very useful that there are some people looking at my work in detail – it's very important'*. Recognising how constructive the examiners' criticism was, the candidate added: *'I'm very happy to have selected them as examiners'* (STEM-C2). For such candidates, acceptance of constructive criticism in itself is one of the attributes expected of a mature professional in any field, including academia.

Observations indicated that not all candidates were equally appreciative, even if accepting of examiners' feedback. For example, one candidate did not overtly agree with some critical comments from the examiners about the structure and coherence of the thesis earlier in the viva, yet towards the end when discussing the possible publication of a book acknowledged that parts of the thesis did not necessarily add anything to the argument and would not be needed in any publication. Another candidate was surprised at the examiners' attention to detail in the thesis, having expected questions to elicit a broader knowledge of the field, such as why the work was important, but agreed afterwards that this had been necessary in relation to awareness of statistical techniques and in preparing them to be a scientist.

Feedback on the thesis, often linked with the formal corrections required before the candidate can pass the examination, may contribute to maintaining thesis quality and standards of presenting research. In making corrections, a candidate may also be increasing their understanding, for example, of statistical or other methodological techniques. Evidence also suggests that some of the feedback provided helps candidates to know what is required of them in a professional capacity, especially in academic life. In these senses, the viva appears to be a learning opportunity for candidates, rather than simply a summative examination. Some feedback to candidates concerned ethical research practice, for example, how to present multiple-authored papers when included in the thesis and the need for clarity about the candidate's precise input.

4.5 Confirming authorship

Several candidates suggested an important purpose of the viva was to ensure no plagiarism had occurred and that the candidate was indeed the author of the thesis. For example, one candidate, even before being asked about the viva's purpose, suggested: *'Actually I think the viva is important in proving that you wrote [the thesis]'* (AHSS-C1). Another candidate, when asked why it was necessary to have a viva, first answered: *'Well, one reason is plagiarism, to check for that'* (AHSS-C3). A third responded: *'Well, really, it's a kind of confirmatory thing. In most cases, they're confirming that you wrote [the thesis] and that you know what you're talking about'* (AHSS-C4).

An AHSS supervisor, having summarised other purposes of the viva, responded:

The bottom line is: did they write it? But you don't actually usually get to asking it that crudely because, unlike certain undergraduates that you get to know who obviously don't write their own essays, you know perfectly well that the PhD student has actually had to do so. (AHSS-D2)

An examiner, on the other hand, suggested: *'The purpose is to check that the student has done the work'* (AHSS-I3). This was supported by two more examiners in different AHSS fields: *'It's a chance to...check that this is the candidate's own work'* (AHSS-I2); and *'To ensure the candidate has undertaken the work'*. This examiner, in a practice-based field, also distinguished between the thesis and viva in describing the latter as *'not the work but a process of checking that the work has happened...The viva is basically a check...the thesis is the main thing'* (AHSS-E6).

The view that one purpose of the viva is to authenticate the thesis was shared by some STEM candidates and examiners: *'Well, firstly to prove that you did it, [that] it was your own work'* (STEM-C4); and *'The purpose of the viva is to make sure you have the understanding, that you have done the work yourself...most importantly, that you've done the work'* (STEM-C5).

Another STEM candidate confirmed the internal examiner would be aware they had actually conducted the research project in a particular overseas location and that it would have been difficult to plagiarise the research:

I suppose they have to check [authorship], but [the internal examiner] has seen me quite a lot [over] four and a half years...and...knows that I really went to [country]...In theory they have to check whether you actually know about your research, but...I've given talks at conferences and [the internal examiner] has seen them and so...in practice they know that I did the research (STEM-C1)

An examiner in the same case questioned the feasibility of plagiarism:

How could you cheat in a science PhD? ...An interesting thought. It would be hard to plagiarise without an examiner picking up that the ideas were already in circulation or had already been published...it might be possible if someone other than the candidate had done the statistical analysis but the candidate said they had done it themselves. You could make up all the results but it would be hard to do convincingly. Why would you do it? Making up the data is harder than collecting it. It does not seem feasible. (STEM-E1)

The examiner also acknowledged the role of the viva in such a situation: *'If you were sceptical about the thesis, the viva gives you the opportunity to probe'* (STEM-E1). The internal examiner in the same case acknowledged the opportunity provided by the viva for further exploration, referring to a different examination where the examiners had been able to follow up concerns that the statistical analysis had been undertaken by someone other than the candidate. This led to the candidate admitting that they had not carried out their own analysis and had not understood the statistical results.

A STEM examiner in a different field, agreeing that plagiarism does occur, made some positive comments about university processes for checking authorship referring to the increasing use of Turnitin and citing a case where they had identified plagiarism by comparing different levels of proficiency in English. In future they expected that, rather than relying on the awareness of examiners to spot plagiarism, there would be increased use of computer software. The view of another STEM examiner was that a principal purpose of the viva was to check authorship: *'It's to be absolutely sure that the person has done the work...'* (STEM-E3). Other STEM examiners, however, similarly to STEM-E1, thought it unlikely or infeasible that the candidate might not be the author of the thesis and were likely to assume it was the candidate's 'own work' (STEM-I2). Respondents did not confuse straightforward plagiarism with the need to acknowledge others' contributions to the thesis.

Interview and observation data concerning the purposes of the viva suggest the complexity involved in the continuum of judgement: examiners are combining their evaluation of the candidate's written work with the viva performance and taking into account a range of factors, including the varying quality of theses and the abilities of candidates to defend their research. Two internal examiners provided different yet complementary summaries of the overall purpose of the viva. One summarised it succinctly as follows: *'It's to assess the suitability of the candidate for the award of the doctorate. To find out whether they really know their subject well and can give a good account of what their thesis is about'* (STEM-I2). The other reflected:

The viva is there to have a real, honest discussion about how things have developed through the research, the difficulties...how did you mature through the process...and how you're able to translate this [to] whatever you find...The viva is not there to memorise everything by heart but to put things in context...you need to be very well aware of your discipline, but you need to...explain to the other person what it's all about. (STEM-I5)

A supervisor with experience as an examiner summarised what they perceived as the purposes of the viva:

To address questions which arise in the course of reading the [thesis], which can be of enormous variety...ranging from factual errors to understanding concepts or handling conceptual or theoretical schema or anything of that kind...But...a... fundamental, underlying thing is...can they speak as well as write, can they write as well as they can speak? So if challenged over some formulation, are they articulate in their response to it? ...Is this person who's just written this...sophisticated tome about such-and-such...do they know what they're talking about? (AHSS-D2)

The responses of STEM-I2, STEM-I5 and AHSS-D2 encapsulate many of the reasons implicit in the rationale for an oral element in the examination, highlighting the dynamic nature of the viva and the added dimension provided by a face-to-face meeting with the candidate. None of the examiners interviewed suggested the viva was irrelevant, even though all assigned greater importance to the thesis. The comments of the final three respondents, while suggesting potentially different priorities among broad subject groups, display agreement about: the need for the candidate to be fully conversant with their subject and able to articulate their knowledge; and the opportunity provided by the viva to test these attributes. Candidates in particular appeared to view the viva in a positive light, agreeing with some of the other actors that it was a rare opportunity to discuss their own research with experts in the field and valuing the 'rite of passage' experienced.

To conclude this section, I include a quotation from an experienced AHSS examiner in the performing arts, who emphasised the informal effects of the viva, accentuating its importance in the process and in academic culture and confirming it as the point where the candidate achieves equality in the academic debate:

The viva is also...the...examiners conversing, via the candidate...therefore...it's part of the academic debate. It's not...written down...but it's important because it's an academic dispute...The viva itself shouldn't be considered formal only, because then it's empty...the two positions [the examiners]...can converse, and the candidate is...part of that conversation...this is a seminar on a very high level, by degree, which is important to the candidate of course...but it's also important [for] the debate that is going in the discipline. It's an internal discussion...but its seeds probably will emerge somewhere else in the work of those three people. This is not a function of the viva as such, but it's a result of it...two academics that haven't met normally discuss the work of a third, in a manner which

is...objective, ideally...normally people forget...this is also a normal academic process, apart from being an examination. (AHSS-E6)

This summary encapsulates the approach that some of the examiners I observed brought to the viva. Even when it was obvious the candidate was going to be asked to undertake corrections to the thesis, examiners engaged in an academic conversation while simultaneously examining the candidate. In several of the cases the discussion became similar to the 'seminar' described by examiner AHSS-E6. This suggests that examiners i) remain fascinated by new developments in their field of research; ii) that their focus is principally on the candidate's research achievements; and iii) that assessment of the candidate's professional and personal skills is implicit, rather than explicit.

Having explored the role of the thesis and the purposes of the viva, I now focus on the relative importance of each element and the relationship between them.

5 The relative importance of the thesis and viva and the relationship between them

The concept of the PhD examination as a process with two interdependent elements and a single outcome (Tinkler and Jackson, 2004) has rarely been questioned or explored in the literature, although Morley et al. (2002), argue for further investigation of the relationship between them. Recent Australian studies, including Lovat et al. 2015, imply the viva is an adjunct to, rather than an integral part of the PhD examination. Data generated by this study, however, suggest a complex and integrated relationship between the two elements, demonstrating a range of perspectives on their relative importance and suggesting the distinction between them is not clear-cut.

5.1 The relative importance of the thesis and viva

When asked about the potential for the viva to change thesis-based judgements most respondents took the view that, unless something particularly unexpected happened (e.g. the rare occurrence of discovering the candidate had cheated), viva proceedings were more likely to influence the amount and nature of corrections than to change whether the candidate passed or failed. Nevertheless, the extent of the viva's influence is shown to be variable and dependent on the quality of the thesis and the candidate's performance.

Several examiners believed the viva would not normally affect the pass/fail judgement, including this external examiner who was categorical about the inability of the viva to change a

judgement already made on the basis of the thesis quality: *'Clearly, you can't make up for a bad thesis by a good viva, so the written PhD...is...more important...both components are vital, but I don't think, if you have a terrible thesis, it's salvageable at the viva'* (AHSS-E1). Another AHSS examiner was equally convinced of the stronger weighting of the thesis and had had no experience of the viva changing the examination outcome:

In my experience, PhD examiners give far greatest weight to the thesis...because that's the formal presentation of three years or more work...The essence of the work is in there, and the viva is confirmatory or clarificatory...I can't think of a case where a viva has changed anything really, in any sort of basic sense. (AHSS-I1)

Two STEM examiners agreed: *'The primary basis of the judgement is on the written material...the viva is useful but the main judgement is formed from reading the thesis'* (STEM-I2) and:

The judgement comes as you are reading the PhD thesis. In most preliminary reports you're asked to anticipate what your judgement's going to be. More often than not it would be *'Minor corrections pending a satisfactory performance in the viva'*. Normally the student would have to behave quite badly or extremely well to change this...Even if they can explain everything satisfactorily and they perform well, you still need to clarify certain things. You come to initial conclusions when reading the thesis and keep these all the way through the viva unless something unexpected happens. The viva consolidates and confirms expectations.' (STEM-E1)

It appears that, for experienced examiners especially, the thesis is likely to predict the final judgement unless the candidate's performance in the viva is either much worse or better than expected. The two STEM responses also emphasise the continuum of judgement.

In another case, the examiners had agreed in advance the viva could not change their minds about the significant amendments required and accordingly the viva did not affect the fundamental thesis changes. One nevertheless acknowledged that on other occasions the viva might play a *'bigger role...for example when the corrections aren't so extensive as they were in this case'* (AHSS-E3), and explained:

We agreed the candidate should pass with minor corrections...If the viva goes well, you might not ask for any corrections. Sometimes [it] can make a difference at the margins. Last week's viva didn't go that well – don't know whether it was a cultural thing or about the candidate's personality. In other circumstances we might have pushed [them] more on some things. You have to work out what your aim is if you push a candidate. It can get quite conflictual and emotional. The viva could take a wrong turn and be hard to get back on track. You can't be a Paxman or a Humphrys. Candidates aren't professional politicians. It's an emotional experience...You try to see what you can get out of them. If they're not prepared to give, you just have to move on...It's better to concentrate on the outcome rather than whether the process of the viva is going along OK. It isn't my show.

Some externals think the viva is an opportunity to grandstand or do an ideological critique of the thesis. That isn't my job. (AHSS-E3)

This experienced examiner illustrated how complex a viva situation can become, showing how, even when examiners are reluctant to push a candidate showing signs of stress, they remain committed to upholding thesis standards. The examiner displayed a reassuring awareness of the need not to interrogate the candidate, or to 'show off' during the viva, one of the criticisms of examiners made in the literature (e.g. Pearce, 2005). During this viva, when the examiners explained to the candidate why they were recommending the particular corrections and the candidate was not fully in agreement with them, a tense atmosphere seemed to develop, which the examiners worked hard to dispel. By the end, the candidate appeared to accept the recommendations and did not voice any concerns during their interview. As an observer, it was interesting that what appeared from the content and tone of the viva to be significant amendments turned into a recommendation for minor corrections, possibly because of the candidate's time constraints.

On the other hand, some respondents gave examples of the viva affecting their judgement, showing there was a balance to be struck between the two components. This internal examiner identified the element of uncertainty present in every viva:

You're always surprised...you form a view and someone comes through the door and you definitely get a...different view...It can go either way...some people are much worse at writing and others are better...and some things you just don't know...you've read [the thesis] but you're not quite sure...you wait to see what the person says. (STEM-I2)

They also confirmed that in the case observed, the viva had shifted the examiners' provisional judgement from minor to major corrections:

[The candidate] did worse in the viva than expected...you can do better or worse. I had a more negative view...at the end of the viva than I did at the beginning. I thought...that there would be a few small changes but it turned out to be more substantial, so...the viva does inform quite strongly what you do next.' (STEM-I2)

The ability of the viva to influence outcomes was evident in another STEM case I observed. While the external acknowledged thesis quality was normally a reliable indicator of the candidate's achievement, they emphasised the importance of the viva in this case for a fair judgement:

The viva is usually confirmatory of the quality of the thesis...I can't think of too many occasions on which what would apparently seem to be a brilliant thesis when you read it turns out to have a...disappointing performance by the candidate... Usually, where you've got an outstanding thesis, the candidate is outstanding...in the viva and...if you've got a relatively weak thesis, then generally the candidate isn't so brilliant either...it's great when it turns out that they're better than their thesis.

Demonstrating how the candidate's performance exceeded their expectations, the examiner continued:

This most recent one...was an example where the candidate turned out to be a lot better than the thesis...that actually shows the value of the viva as opposed to...a situation...where people decide...that the viva is not worth having. I...don't agree with that at all because...[the] viva that you witnessed was a good example of a relatively poor thesis, where we were...confirmed, ultimately, in our view that the student should pass and be awarded the degree, based on their...good performance in the viva. (STEM-E5)

In another case an internal examiner, when asked about the possibility of a candidate who had written a high quality thesis not being able to sustain the same level of excellence in the viva, agreeing with STEM-E5, suggested this was rare, yet confirmed the value of the viva as an opportunity for the candidate to provide further information:

I've never been in a viva where a person who has written a brilliant thesis hasn't been able to say anything about it. [The viva] allows the student to expand on or clarify points not in the thesis. Even if they haven't explained something adequately in the thesis they have an opportunity to add this in the viva. (STEM-I1)

A convenor of another viva I observed maintained the viva could make a difference in borderline cases, suggesting that it would be hard to fail if the candidate had written a good thesis but if they were borderline or somewhere between minor and major corrections, the viva could *'push them up'* (AHSS-IC1). They also suggested the viva would only diminish the outcome if the candidate was *'desperately weak'*.

On the other hand, a supervisor and examiner of many years' experience, suggested it would be unusual for a candidate to improve on the impression gained from the thesis, whereas the viva performance of some candidates was poorer than might have been expected from the quality of their thesis: *'I've...rarely had anyone who has surprised me in a positive direction in the viva. I've had a few where I've been rather surprised that the quality of the written product didn't quite seem to be borne out in the viva'* (AHSS-S1). A STEM examiner argued that *'an absolutely terrible viva could overrule a good thesis'* (STEM-E3), leaving open the option of the viva overturning a prior judgement based on the thesis, whereas other respondents suggested it was more likely for a borderline thesis to be *'rescued'* by a successful viva examination than for a candidate who had submitted a high quality thesis to give a disappointing viva performance. For example:

There are some people who appear to be very sophisticated in argument, who aren't when you talk to them, who are much more tentative...uncertain, and reveal areas of lack of knowledge that you didn't think they would have. (ECON-EX1).

This examiner nevertheless confirmed that such situations had not actually changed their view or the overall judgement.

Another response emphasised the relative importance of the viva for fine-tuning judgements:

The viva is really important because it does allow you to probe and test and evaluate...It's not going to sway you between...passing without corrections and a fail, but certainly between the various levels on the ladder...the viva can play a determining function. (AHSS-I4).

These responses capture the collective view of examiners concerning the primary role of the thesis and the confirmatory, consolidating role of the viva. The comments of the next examiner, on the other hand, contain some potentially conflicting views, first stating that both elements are necessary, yet suggesting the viva could be dispensed with, and finally confirming its value in testing skills not apparent from written work:

You definitely need both, although in Australia they get away without having vivas...The system hasn't fallen apart without the viva. That suggests the viva is not that important. The rest of the person's scientific output is largely judged on their written work. The thesis is absolutely critical. The viva is possibly dispensable, if either part is. But, the viva gives you a chance to explore some of the issues. It supports quite a traditional approach to teaching: testing, debating and other scientific skills that a scientist should have, which are not gained from just writing. (STEM-E1)

Although ambivalent about the viva, this examiner acknowledged it facilitated further exploration of the candidate's attributes and importantly tested some personal and professional skills, albeit implicitly.

Making another international comparison, a non-case related examiner referenced an experience acting as external examiner for a German PhD viva where the viva is by way of a public defence and explicit weightings were given for different elements of the examination process, as follows: thesis 60%; pre-viva presentation 20%; viva 20%. In addition to providing overt weightings for the three components, this also raises the question of whether candidates should be offered the opportunity to present their findings at the beginning of the viva, a suggestion made by one of the candidates I observed, but declined by their examiners.

Another examiner, in a case where the candidate had submitted for a PhD by published work involving practice, suggested that in such cases vivas were even more important in the overall assessment:

That's why the viva here is...more crucial than in other areas. If you write a thesis, the viva in itself is much less important...[in] a normal thesis...the two or three examiners...are so well-versed with both the viva and the background that the viva in itself is almost a

performance. But in this case it wasn't, because there are two...distinct types of material and the relationship between them is what we are judging and also...the shortness of time, etc. So the viva here was more crucial, will always be more crucial, and it had to satisfy us totally. It also had to satisfy the candidate that they know what's wrong and what needs to be corrected and therefore, it's not a question of...ticking boxes but the broad arguments had to be made, heard, understood, reacted to, questions asked etc., so this is not a performance, this is the real thing.' (AHSS-E6)

The examiner is suggesting that in cases involving published work with analytical commentaries, particularly if they include assessing artefacts, much hangs on the viva. Examiners and candidates regard it as a serious opportunity for making arguments to support the thesis and related work, exchanging views and explaining perspectives and corrections.

Analysis of the assessment process suggests examiners are aware they use the viva to fine-tune their judgements, especially in cases where the thesis raises concerns. Their responses suggest considerable but not full agreement about the relative importance of each element, as reflected in the approximate percentages summarised in Table 5.1. While not asked specifically to do so by the researcher, 11 respondents provided approximate percentage weightings for the relative contributions of the thesis and viva that they would generally give to the overall examination judgement (not necessarily specific to the case observed). The highest percentage allocated to the thesis was 95%, and the lowest 70%. The average of all the approximate percentage weightings was: thesis 79.5%; viva 20.5%. Full details of the approximate weightings suggested by respondents are provided in Table 5.1, by role.

Internal examiners		External examiners		Candidates		Supervisors	
ID	Percentage weighting thesis:viva	ID	Percentage weighting thesis:viva	ID	Percentage weighting thesis:viva	ID	Percentage weighting thesis:viva
STEM-I2	75:25	STEM-E2	70:30	AHSS-C3	80:20	AHSS-D1	80-90:20-10
STEM-I4	95:05	STEM-E3	75:25	STEM-C4	75:25		
AHSS-I4	80:20	AHSS-E3	90-80:10-20				
AHSS-I3	70:30						
STEM-I1	85:15						

Table 5.1: Approximate percentage weightings showing relative importance of the thesis and viva in the overall judgement

While it is not possible to draw any firm conclusions or perform a statistical analysis on such a small sample, it is of interest that most of the 11 respondents put the percentage weighting at around 80 for the thesis and 20 for the viva, reinforcing the importance of the thesis in the overall judgement. The variation among the respondents may reflect individuals' approaches to and experience of the PhD examination, including differences in thesis quality. In suggesting a numerical range within the approximate percentage, the AHSS external examiner and supervisor (Table 5.1, columns 2 and 4) are potentially implying that the relative importance of the viva differs from one examination to another. Another interesting feature is that almost equal numbers of individuals in both subject groups felt able to suggest approximate percentage weightings, with divergent views within subject groups, as well as between them, as illustrated by the internal examiners' weightings in column 1.

Candidates' views

In most cases, candidates were in agreement with examiners in assigning primacy to the thesis, for example: *'...I think my [thesis] was the most important, as a kind of synthesis of my work and my academic research, so I think on balance the supporting statement...is the most important'* (AHSS-C5); *'Well I think the thesis is more important because without the thesis you don't do the viva, but I am pleased to have done the viva'* (STEM-C1); and *'I think [the examiners] put more [weight] on the thesis'* (STEM-C2). However, candidates also considered the viva to be significant in the overall judgement, some even suggesting it had equal weight with the thesis: *'It feels like the thesis is the most important thing when you're doing it but now I suspect that maybe it isn't, maybe the viva is just as important. Certainly people who have spoken to me who are PhD examiners have suggested that the viva is just as important.'* (AHSS-C1). Another candidate supported this view and alluded to the personal and professional skills tested in the viva:

It is essential to have the viva because...you need to show that you're eloquent, showing what you've done and what you're thinking about in writing, but actually to do ad hoc problem solving, to maybe take the ideas of the thesis even further and to really see almost how the mind of the scientist works, you can really only do that by talking to them...they're both essential. (STEM-C3)

A STEM candidate who gave more weighting to the thesis also highlighted the opportunity provided by the viva to expand on and defend the written work:

The thesis is more important, but...the way you speak in the viva is...very important because, if you manage to defend the things that you've done...and the way you've done it, they might change their mind if they wanted to...change something in your thesis. But

if you can defend it well and also...get [your] point of view...across...that's why the viva itself is also important. [STEM-C5]

An AHSS examiner who claimed to have entered the viva with an open mind about the final outcome, relied above all on the student's viva performance to confirm thesis-based judgements, which they considered 'very important' (AHSS-I4).

Another examiner's atypical view suggested the viva was more important even than the thesis:

The viva is very important...far more important than the thesis because, [in]...the viva, you can actually see [what] they really know...The thesis gives you an idea of the type of research project, how good or bad it was, but about the candidate, you confirm – if the thesis is well-written and you have a very good viva, you just confirm that [they] did very well, but if [they're] a lousy candidate and a very good thesis, most probably the supervisor did a lot of the writing. (STEM-I5)

This quotation raises many pertinent issues concerning the viva, including the potential variability of the examiners' approach and the extent of supervisor intervention. The examiner seems to imply that it is not possible to come to a final judgement on the examination without meeting the candidate. The last three examiners viewed the viva as absolutely the final point of judgement, irrespective of prior conclusions based on the thesis. They also illustrated how the continuum works in practice, the viva forming the final stage of an integrated judgement.

Most of the actors involved in the process involved in this study considered that PhD examination judgements are better informed as a result of the viva which, it could be argued, provides an opportunity for examiners to gain a rounded view of the candidate by examining their professional and personal skills as well as deepening their understanding of the written work. Responses also suggested, as in some literature, that examiners usually want candidates to pass and adopt an encouraging, rather than confrontational approach. Two examiners' responses suggested they were interested in helping candidates to do their best in the viva. One referred to the skills needed by examiners to draw out information from more reticent candidates, while the other, having suggested it was important for the candidate to include in the thesis a general discussion to demonstrate the significance of their research in the wider field, confirmed that exploring the candidate's ability to '*bring it back to those big questions that they pose in the introduction...tends to be something that's nice to explore in the viva*'. They explained that if evidence that the candidate had understood the significance of their research was missing from the thesis, it could be evaluated in the viva. If it emerged '*that the understanding is missing...that's more problematic and...more in the direction of major*

corrections' (STEM-E4). Both examples show examiners wishing to facilitate a fair outcome by using the viva to the candidate's advantage.

According to the majority of my respondents, the thesis predominates, yet thesis-based judgement is contingent on a satisfactory viva. One or two examiners, however, similarly to a minority of candidates, attached greater importance to the impact of the viva on the final outcome. No obvious differences were apparent between STEM and AHSS disciplines concerning the relative importance of each component.

5.2 The relationship between the thesis and viva

In explaining how judgements were arrived at, some of those interviewed for this study suggested an interdependent relationship, with greater emphasis on the viva if the thesis had raised concerns, for example:

There are situations where a rather 'wobbly' thesis is well-defended at the viva and so a candidate convinces the examiners [that] actually they're much better than the written thesis gave the impression that they were...so the examiners are more sympathetic to them, having heard the defence of it at the viva. (AHSS-E1).

The role of the thesis includes raising questions for the viva, and the difference between minor and major corrections may rest on the candidate's answers. For example, this examiner, having confirmed their judgement was principally made on the basis of the thesis, described some of the characteristics of the thesis that led to questions in the viva and illustrated how answers enabled examiners to refine and consolidate judgements:

In this...example, the first chapter was published but [the candidate] didn't say so in the written work...we actually knew that before, but that came out in the viva. [The candidate's] statistical analysis was inadequate but it wasn't totally clear why that was...what came out pretty strongly was that [the candidate] literally didn't know enough. So there are things you have to elicit in the viva. (STEM-I2)

In this case, the thesis raised numerous questions. I was able to observe the examiners' concerns emerge during the questioning and it was clear the viva played a significant part in the judgement, changing the outcome from potentially minor corrections to major corrections.

The external examiner below illustrates how the viva of a practice-based candidate was used in conjunction with the analysis of the submission to explore some methodological questions and to encourage the candidate to review data sources:

[This candidate's] work...is very interesting. It's very unusual. There are moments of brilliance about it, which were borne out in the viva. What was less clear, from reading...is

if they...actually managed to control the sources and the range of sources. I was unhappy...with the range of sources used, and thought there were important methodological arguments that the viva had to clarify. If they were unaware of or unwilling to use those [sources]...that was a problem for me. (AHSS-E6)

In this case, the examiners' judgement was three-dimensional, taking account of the artefacts produced by the candidate, the written submission and the oral examination. The examiner's comments suggest all three dimensions were interdependent and contributed to the overall judgement: the impressive technical work, the need for enhancements to the accompanying submission and the need for the candidate to provide clarification in the viva. The outcome was a requirement for substantive yet 'minor' corrections to the thesis, the examiners jointly recommending that the candidate should re-write the introduction and add to the literature review.

These two cases emphasise the integrated nature of the relationship between the thesis and viva (and in the second example the practical element), a view articulated by others, for example: *'The judgement about the written thesis has to affect the viva'* (STEM-IC3); *'It's not really possible to envisage a PhD exam without a viva'* (BIO-EX2); *'I think it would be impossible to imagine a PhD examination process which didn't include a viva...you have to have the person there, to defend what they're saying'*; and:

The weighting question suggests that one can separate the two out, and I don't think you can...the thesis is the necessary but not sufficient...passport for completion...you have to have a defensible thesis and defend it...if you don't have a defensible thesis, you can't defend it...the viva is important, but I [can't]...see it apart from the thesis' (AHSS-I2)

Two examiners, perhaps surprisingly given the dominance of the thesis in the examination, suggested that it was not until some way through the viva that they came to a final decision about the examination outcome. The first made a general distinction between cases where the thesis is sound and those where as a result of reading the thesis the examiners have residual concerns:

If you...feel reasonably comfortable and the viva is almost...something you're...going through because you've got to...but...from...what you've seen and read ...the student is going to pass...because it's a perfectly sound piece of work and the viva is just the person coming along to...confirm that...that's rather different to where you've got serious concerns and the student's got to...actually defend the thesis as an...integrated piece of work...In those instances where I have had doubts...you come with a set of quite serious questions that the student has to respond to satisfactorily...and so you're probably waiting till quite late on in the viva, three-quarters of the way through, I would say, which would be the kind of tipping point. (AHSS-I4)

This examiner was clear that in potentially borderline cases the viva makes a significant contribution to the examiners' final judgement. The next examiner draws on the case observed to illustrate how the candidate's performance in the viva gradually convinced the examiners they should award the PhD:

We had to go quite far through the viva before I was convinced about the candidate...we needed to ascertain that the student understood the research and results they were presenting, and there were times...when the student was found wanting, but never hugely so...[they] convinced us more and more...that...[they] had done this work...pretty much in isolation, quite often having to seek out additional support...[They were] able to demonstrate... those attributes that I was looking for, but...there was certainly doubt...before we met the candidate...but [they] did inspire more confidence as the viva went on.' (STEM-E5)

In the latter case, the examiner's response is entirely in alignment with what I observed during the viva and the way in which it was managed by the examiners. AHSS-I4 and STEM-E5 gave practical examples of how the continuum of judgement works and substantiated other responses by suggesting that the viva has more significance in cases where, having read the thesis, the examiners have residual concerns. They are then seeking a good performance by the candidate in the viva to explain the weaker elements of the thesis and allay some of those concerns. In order to satisfy the examiners in this respect, candidates need to display a range of personal and professional skills, not necessarily evident from the thesis. On the other hand, as Murray (2009:19) suggests, *'in the case of good or outstanding theses the oral examination matters less, though it cannot be said that this is always the case'*, the ambiguity arising from the extent to which examiners treat the oral examination as integral. In my study, however, all examiners treated it as a fundamental part of the examination.

One candidate who was progressing to an academic role, having already expressed the view that the viva was an *'essential'* part of the examination process, referred directly to some of the personal and professional skills that can be evaluated in the viva:

You can write whatever you want [in the thesis] and it will be checked...It's not that you write something one day and then never look at it again before you submit it...[the thesis and viva are] both essential but they're testing two different things. One is being eloquent on paper which is very important for writing research papers and grants and the other is actually how you're thinking and that you're eloquent in explaining what you're thinking with your voice, because that's essential too...you need to be presenting your data, you need to sit in grant reviews and talk about what you want to do and what this all means, so I think they're both essential. (STEM-C3)

It could be argued that the thesis and viva taken together as a single examination should enable PhD examiners both to evaluate a candidate's ability to carry out independent research

in the field to a particular standard that takes them over the threshold level required for a PhD, and simultaneously to assess the personal and professional skills they can demonstrate after experiencing around 4 years of full-time (or the part-time equivalent) doctoral study under supervision in a structured research environment. However, this perspective is not necessarily widely shared and data emerging from this study suggest that examiners' priorities are connected with the former purpose of the assessment rather than the latter, although both interviews and observations demonstrate that some personal and professional skills are necessarily assessed throughout.

6 Conclusion

When the doctorate was first introduced in the 12th century in Paris, the degree was awarded after a series of oral examinations or defences during which the candidate was required to be proficient in at least two languages (Noble, 1994). Since then, however, and for many years the thesis has assumed prime importance in the judgement process. Unsurprisingly, the widely held view of the centrality of the thesis to the PhD examination is supported by this study. This is particularly evident in its influence on the examiners' approach to the oral examination, with a high quality thesis raising expectations for the viva and a 'wobbly' (AHSS-E1) thesis causing concerns that the candidate may disperse in the viva. One of the most striking impressions emerging from examiners, candidates and supervisors is the way in which the two elements of the examination are intrinsically linked in a continuum of judgement that begins with initial evaluation of the thesis and finishes at the end of the viva. Although three stages of this continuum are suggested (Figure 5.1), the reality may be more complex, depending on the quality of the thesis, information the examiners discover in the viva, the performance of the candidate and their future career. The relationship between the thesis and the viva is therefore multifaceted and interdependent, even if not equal. The data collectively suggest that while most examiners' views do not change radically as a result of the viva, nevertheless it fulfils important purposes for both examiners and candidates and there is considerable agreement between them on what those purposes are. Some examiners go as far as to suggest it is impossible to imagine a PhD examination without the viva, suggesting that to be fair to the candidate it is necessary to meet them face to face and for them to have an opportunity to defend their work, a view shared by candidates. Murray's summary suggests the contradictions inherent in the relationship:

There is some evidence...that the viva is itself part of the assessment, although there is also evidence that a student's performance at the viva will not change an examiner's mind. This means that the oral examination still constitutes the final assessment and still

requires an excellent performance from the student. It is, after all, compulsory...The crucial – and perhaps initially alarming – inference here is that the oral examination may – or may not – decide the outcome of the examination of the thesis. (Murray, 2009:18-19)

What this quotation does not make clear, however, is that although in the UK the viva is compulsory, the relative importance of the two components is variable and depends on several factors; one supervisor even suggested: *'...the relative weighting of the two components...varies dramatically across different students'* (STEM-S4), making the point that some students write well but do not perform at a high level under pressure in a situation where two examiners are asking challenging questions.

One of the clearest outcomes from this study regarding the thesis-viva relationship, however, is that the viva can be most influential in cases where the thesis raises concerns. Additionally, some of its purposes, for example providing the opportunity for examiners to confirm and modify their judgements and to achieve a deeper understanding of the candidate's research, are significant, irrespective of thesis quality. One of the implicit themes emerging from this chapter is that to do well in the PhD examination, candidates must display a range of professional and personal skills, including the ability to communicate effectively. Some skills are apparent in both thesis and viva; others, such as some of those articulated by candidates, are best tested and demonstrated in the viva. The importance of these skills, however, was not articulated in a way that suggested that assessing them was a main purpose of the examination.

The dynamic nature of the relationship between the thesis and viva is a core element of the process through which examiners make their assessment of doctoral candidates and, although most examiners are clear that the candidate stands or falls by the content of the thesis, they acknowledge that the viva plays a crucial role in the process and therefore in the continuum of judgement. Within the viva, the convenor has a quality assurance role that goes some way to moderating the process.

Chapter 6: Attributes sought by examiners in the thesis and the candidate

The scholar is expected...to think for himself and to develop an independent and critical mastery of whatever subject he may have at hand...he is expected to know all that has been learned up to his day in respect to this subject, and...to know and digest the literature pertaining to it...by personal contact with the original facts, by seeing for himself and by examining them by the best methods and in the most thorough way, he is expected to learn something not yet known, and thus to add to the sum of human knowledge. (Thwing, 1928:58)

1 Introduction

The requirement for candidates to submit work that is original or makes a contribution to knowledge has long been a defining criterion for the award of the PhD. Although outdated in its use of gender-specific language, the above summary of the technical expectations to be fulfilled by doctoral candidates, especially to *'add to the sum of human knowledge'* remains pertinent. A more recent expectation, though in most subjects not a requirement, is that some or all of the work should be considered by peers to be worthy of publication. When the PhD was an *'elite'* degree intended principally for those entering academic careers, it was tacitly assumed that the PhD *'standard'* was understood by the academic community engaged in its examination. It was also accepted that the standard was based on academic judgement, using evidence that was often implicit (Lovitts, 2007). More recently, the rapid expansion of doctoral education together with the introduction of explicit standards for all degrees has increased the significance of other characteristics overtly sought by examiners. This shift was reflected in some of the more professional and personal attributes referenced by examiners in this study.

The inclusion of doctoral learning outcomes in European and UK-wide qualifications frameworks (Bologna Working Group on Qualifications Frameworks, 2005; QAA, 2008) and the increasing diversity of candidates and employment destinations have caused universities to attempt to develop more explicit doctoral assessment criteria. Some include characteristics transferable to other contexts, such as: *'the exercise of critical judgement'* or *'intellectual coherence'*. Others include expectations that candidates will develop more specific professional skills such as leadership. Nyquist (2002:19) acknowledged that *'Since by definition, the PhD represents the individualised development of a researcher/professional, it is difficult to think of a generic set of competencies for doctoral recipients'*. Given the different cultures and expectations of academic disciplines and the idiographic nature of PhD

assessment, introducing more explicit assessment criteria is a challenge, made more complex by evolving PhD structures and purposes.

In this study one respondent showed that individual judgements are not easily accommodated by more structured assessment by reflecting on the lack of transferability to the doctoral examination of the criterion referencing prevalent in taught degrees:

One of the curiosities of examining doctorates is that...whatever university regulations say, there's no marking scheme. You're actually bringing [your own criteria] to bear on a particular piece of work and...hoping that others are doing likewise. (AHSS-I1).

It is unsurprising therefore that the attributes PhD examiners seek when assessing candidates and their work are the subject of considerable speculation, exacerbated by the lack of transparency in the way judgements are made and by the discipline-specific approaches demanded by the research. Despite these difficulties, in this study respondents were able to articulate clearly a range of attributes sought by examiners.

Institutional assessment criteria are intended both as a guide for candidates and examiners and to demonstrate the threshold standards for the award. Such criteria are of necessity set at a more general level than subject-specific expectations and include requirements for 'originality' or a 'contribution to knowledge' (Lovitts, 2007; Clarke and Lunt, 2014). Both originality and publishability commonly feature in the formal guidance for doctoral assessment provided by universities, for example: 'originality and creativity', 'significant and substantial contribution' and 'of a quality to satisfy peer review and merit publication'. Criteria may include both academic and professional attributes, such as the UK-wide learning outcomes enshrined in the doctoral qualification descriptor in the FHEQ (Table 2.1), which are intended to apply to any doctoral candidate at the point of graduation. Appendix 6, section 2, contains examples of doctoral assessment criteria from several universities, some of which are explicitly mapped to the doctoral qualification descriptor.

Only one examiner in the current study chose to link the expectation of originality to fulfilling the criteria set by the candidate's university, rather than offering their own interpretation. While the existence of national and institutional criteria is important, data in this study suggest that examiners interpret these criteria according to the implicit standards for a PhD in their field.

This chapter includes evidence from 41 (95%) of my 43 interviewees and five focus group members, as follows: eight external and nine internal examiners; eight candidates; five supervisors, six non case-related examiners, two convenors and three focus group members.

The chapter explores the research question: *'What attributes do examiners seek in the thesis and the candidate?'*. Unsurprisingly, the attribute of 'originality' (section 2) featured prominently, though examiners held a range of views about its meaning, its importance and, crucially, its practical application. Interestingly, candidates considered that originality should be central to their thesis. Section 3 presents findings on the attribute of 'publishability', often mentioned by examiners, some suggesting it was an indicator of the candidate's exceptionalism. For others, however, the expectation of publication during the PhD was considered challenging for all but outstanding candidates, especially in professional fields. After originality and publishability, both principally judged through the thesis, examiners described a range of characteristics or attributes commonly sought in the PhD examination. Evidence of research competence, including appropriate methodology and data analysis skills, showed some discipline-related differences of approach and are explored in section 4. More general characteristics concerning research integrity are discussed in section 5. While examiners situated these attributes in a research-specific context, some are equally important as personal characteristics. Attributes displaying intellectual rigour are explored in section 6. As this chapter demonstrates, examiners cited research-specific, professional and personal characteristics among the attributes they sought in PhD candidates. Responses demonstrated a strong relationship between the continuum of judgement and the thesis-viva-nexus, a term devised to represent the two facets of the candidate displayed in the PhD. What separates the PhD from all other qualifications in higher education is its role in the discovery of new knowledge. This primary requirement for originality is explored below.

2 Originality (or, contribution to knowledge)

The single characteristic most frequently associated with research degrees and their outcomes is 'originality'. Subsumed in this concept is the understanding that PhD candidates will to a greater or lesser extent contribute to existing knowledge in their field. While most examiners consulted in this study did not find the concept of originality problematic, the extent of its importance varied and multiple interpretations were apparent. For some, the primary need for originality remained so intrinsic that other attributes were not considered until that had been established, whereas for others the originality of the work appeared less critical than a range

of other attributes evident in the thesis and candidate. A few examiners even suggested that the concept of originality was problematic or irrelevant.

Typically, institutional criteria concerning the need for the candidate to produce some form of new work are phrased as a requirement for 'originality', 'creativity', 'the creation and interpretation of (new) knowledge', or 'a distinct and significant contribution to the subject'. One examiner supported the use of 'a new contribution to knowledge' as a criterion that might be used in university regulations as an alternative to originality to differentiate between average and exceptional PhD candidates. They also acknowledged other attributes may be equally important:

It is not necessarily about originality. Universities' expectations of PhD candidates usually include 'a new contribution to knowledge' or words to that effect. This is the definition...that I mostly come across as an external examiner. How do you define originality? Maybe you could describe a new contribution [to knowledge] as original.' (STEM-E1)

Institutional criteria suggest that the interpretation of the commonly used terms in universities' doctoral assessment criteria is anything but straightforward.

Many examiners in this study found it challenging to define the abstract concept of originality. Given the expectation of knowledge creation associated with the PhD, this was surprising. In a few cases, diverse interpretations were explained by subject-specific differences.

The challenges inherent in interpreting originality were encapsulated by this non case-related examiner who, after summarising a range of attributes sought in PhD candidates, added: '*Originality is...important but...there are different approaches to how one might define[it]*' (ECON-Ex1). A convenor also suggested multiple definitions: '*Interpretations and expectations of originality vary...[examiners] require the work to have made a contribution to the field. Originality is value-loaded...[There is] no clear definition*' (STEM-IC3).

This section explores originality as: a problematic concept; a priority; one of a range of significant attributes; and as interpreted by examiners and candidates. It is unclear if different examiner perspectives emanated mainly from subject-related expectations or evolved in response to individual examiners' experiences, or both. However, as suggested by Lovitts (2007), attributes sought by examiners including evidence of new knowledge may often be implicit until they are asked to reveal and prioritise them.

2.1 Originality as a problematic concept

For four examiners, originality was: interesting; problematic; not a particularly helpful concept; or a matter of interpretation, for example, '*Originality? An interesting issue; I struggle with this sometimes*' (STEM-E1). A non-case related examiner from a different STEM background was sceptical about originality in the literal sense and suggested it could be demonstrated in different forms, relative to what had gone before, especially in some subjects:

Words like 'original' are bandied around...they're not very helpful because...many of us develop our own ideas which to us are original when, to the subject, they're not...it's...fine that someone is original in their thought...without necessarily having been original to the subject...and that's why [in subjects] that are quite discursive...most of these ideas have been around before...yet no-one would say you shouldn't award any PhDs [in certain subjects] because nobody can be original in the wider sense...[whether something's original] is a dead subject as far as I'm concerned – it doesn't move us on. (EARTH-Ex1)

That an idea may not be original to the subject but to the individual presenting it, as long as there is something new in the way the research is approached is challenging in the context of the PhD, an award whose *raison d'être* is the creation of knowledge. However, the candidate's awareness of existing knowledge in the field, through a literature review or equivalent, provides an opportunity for candidates to situate their research in the wider field and to demonstrate their own contribution.

The next examiner, from a different field, supported EARTH-Ex1's views, expressing the point through emphasising the problems of interpretation:

Originality is a tricky one to define...I don't find it that helpful a term...in my field...the chances are you'll say something new...just by the very act of being in there because the combination of you...and your relationship to the people you're working with is new and different and no-one else will have done that before...I'm not that interested in originality. (AHSS-I2)

These comments illustrated how conducting research in a particular field using discipline-specific research methods can influence definitions. A third examiner's response, again in a different field, also demonstrated the subject context:

Originality I'm not that bothered about...if what they've done is...taken methods that have been used well on one species and then they've applied them to their species and come up with new findings...that's absolutely fine. (STEM-E4)

Rather than suggesting that work could demonstrate originality by applying new methodology to an existing problem (EARTH-Ex1), STEM-E4 argued that using methods proven to be

successful with one species (or in one situation) could be applied to another to produce new knowledge.

Two more respondents displayed contrasting views about originality while finding it problematic. The first suggested that originality was more of a preoccupation for students than for them and described different approaches to defining originality:

Students sometimes get frightened by the word 'original'. No work is original...in one interpretation, we all stand on the shoulders of giants...I tell my students that...broadly...it's less important...published work is supposed to be original – what counts as original? Looking at things in a different way?...That can be looking at the same things in a different way, whether that's updating or from a different perspective, or answering new questions...Originality...strictly interpreted, is overrated...things like originality...are just so elastic. (ECON-Ex2)

The supervisor, however, suggested that the requirement for a candidate to demonstrate the attribute of originality or a contribution to the field may be so embedded in examiners' judgements that it is often taken as read (Lovitts, 2007). When asked directly about originality or a novel contribution, the supervisor explained:

You almost take [originality] for granted. But I...agree that...in not just PhD theses but generally how we judge science and scientists, the novelty and the originality [are very important]...people almost...tacitly assume that that's a major criterion and don't spell it out. (STEM-S2)

Irrespective of how it is described, the concept of adding to knowledge in the field, whether 'incremental' or 'ground-breaking', is intrinsic to academic research and therefore to the PhD. Arguably it remains the most important criterion by which candidates are judged. The comments of the six respondents above introduce the complexities surrounding interpretations of originality and whether it remains a priority in assessment of the PhD, as explored below.

2.2 Originality as a priority

By contrast with the respondents in section 2.1, when questioned about the attributes they sought, three AHSS examiners immediately referred to the requirement for originality or a contribution to knowledge. This suggested it was a priority for all three, yet with varied interpretations. For example: '*The main thing is whether it's an original contribution*' (AHSS-13);

The candidate might not be an original thinker but may be capable of producing work that has enough originality to meet the criteria, for example, knowledge put together in a way that hasn't been done before' (AHSS-E3); and

If the work is not original, and that's...a judgement which is based on a lot of your experience...if you've seen quite a lot of this before, if the argument is repetition of other arguments...it doesn't satisfy PhD requirements...it's not innovative and original and therefore it's not really a PhD. (AHSS-E6)

A convenor agreed: *'The key question is: "Is this a contribution to knowledge?" You can forgive all sorts of things if what's presented has made a significant contribution'* (AHSS-IC1). For these four respondents, originality was a priority and, if present, would outweigh weaknesses in other areas.

Examiners who prioritised originality over other characteristics were in the minority, most indicating it was one of a range of significant attributes sought. This approach is reflected in many institutional regulations, where 'originality' and/or a 'contribution to knowledge' may be mentioned first but are not prioritised.

2.3 Originality as one of a range of significant attributes

Some respondents gave originality equal status with other significant characteristics. For example, an examiner in a technical field suggested that a candidate would need to demonstrate three principal characteristics: *'originality, significance, technical competence'* (COMP-Ex1), giving originality parity with significance and technical ability:

I'm willing to let them off a bit in one of those categories so long as the other two are there...I think there is a balancing act...they're...like a triangle and you...don't have to put your weight equally on all [of them], but you're looking for those three things... (COMP-Ex1)

The examiner later implied that if all three attributes appeared to be present after reading the thesis even if not in equal measure, the viva allowed the candidate to prove they had done enough to be awarded a PhD: *'And if you've got...an ink tick in two of them and a pencil tick in the other one, then there's a PhD in there somewhere and we're just looking at how to get it out into the daylight'* (COMP-Ex1). This examiner took a pragmatic approach when judging whether the characteristics they were seeking had been met, implying that all were important, but that to pass, the candidate need not be equally impressive across all three areas.

An AHSS examiner included innovation and originality in a wider range of attributes they were expecting to find in PhD candidates and their work: *'Innovation and originality in terms of moving the debate of the discipline forward, adding something that wasn't there, developing a set of arguments in a way that is new'* (AHSS-E6).

2.4 Differences of interpretation

Interpretations of originality were contradictory, some influenced by how the term was conceptualised by different subject groups (e.g. AHSS or STEM or AHSS), yet collectively there was some agreement when differentiating between average and exceptional candidates.

For some examiners, the originality or newness emerged organically from the work. For example:

There has to be...a body of knowledge or information that comes out of the thesis...which is new and which adds to the field...some sort of evidence of originality of thought or experiment. (STEM-I3).

An AHSS examiner also linked the attribute of originality to the candidate's ability to reflect and interpret the topic:

What highlights whether somebody's made an original contribution is their ability to reflect on what they've done so they're not just regurgitating what everybody else has done or said, but there is an element within it of their own thought process...and their own conclusions being drawn from the data they've collected...whether it's secondary or primary data. (AHSS-I3).

These examiners' comments aligned with the comments of EARTH-Ex1 (section 2.1) in that both were seeking evidence of original, independent thought in the candidate even if it concerned a long-standing area of research. They also illustrated how the nature of the field influences judgements about attributes or characteristics, one referencing laboratory experiments, the other primary and secondary sources.

Some separated the concepts of 'originality' and 'a contribution' to differentiate candidates who demonstrate outstanding achievement:

It's very important that people make a substantive contribution, but my experience on that one would be that what I consider to be a really original piece of research that's really going to change my mind about [the field] is something I don't come across very often...someone who is really exceptional. (STEM-I2).

This examiner suggested they were likely to judge most PhD research as *'incremental'*, contributing to knowledge but not stunningly original. Similarly, the response above from examiner COMP-Ex1 suggested that in their field also, 'originality' and 'significance' represent different levels of achievement. This AHSS examiner's definition of a contribution to knowledge, on the other hand, is opposite to the STEM definitions and suggests that originality is demonstrated by the generation of new data and a novel approach and that what makes it significant, or not, is the candidate's intellectual rigour in interpreting it:

There's originality in...that there's new data...and...by having investigated something, the candidate has reported...something...novel, but the question, ...is whether they reported it in a way which stands up, such that it does relate to the research questions that they've set...[and]...adds some new insight to the field they're working in...contribution to...knowledge depends what wording you use. (AHSS-I1)

Here, the candidate's 'originality' was not in question, since the data were in some way new. For the examiner it was the approach to the empirical research and the candidate's method of analysing it that may lead to new insight or contribution to knowledge in the relevant field.

Contrasting interpretations were also evident within the two broad subject groups. For example, two STEM respondents provided alternative interpretations of 'novel'. In expanding on their response concerning originality (in 2.1 above), examiner STEM-E4 added that they were seeking:

A contribution to knowledge, not really originality...Well I guess it is originality in that it's not a...straight replication of something...already... published [but]...to me, originality is coming up with something completely novel. (STEM-E4).

While these articulated the examiner's interpretation of originality as something 'completely novel' (or perhaps 'truly original'), a supervisor in a different STEM field conversely conceptualised a 'novel' contribution as an attribute that would routinely be demonstrated by PhD candidates, rather than describing work that would stand out as highly original:

I've never seen a PhD project that wasn't novel...you wouldn't, as a scientist, embark on a project unless it was novel, so the mere replication of something we already know wouldn't be considered a useful project. (STEM-S2).

The supervisor continued:

Originality is an interesting point...what you'd ideally like to see is a project that's... original and where there's...evidence from the viva that this originality is actually [of] the PhD student's making...beyond just being novel...it is nice as a scientist to see some sparkling things...that are really original...not just novel in the fact that no-one's done the exact thing before, but which really break new ground... (STEM-S2)

This supervisor, also an experienced examiner, interpreted 'novelty' in the opposite way from examiner STEM-E4, whose definition of 'novel' implied rarity and appeared to be the equivalent of the supervisor's definition of ground-breaking research with the student leading the work. They implied that the viva enabled examiners to probe the background to the research and the extent of the student's initiative, creativity and 'exceptionality'. These contrasting perspectives imply that terms commonly used to describe originality – in this case novelty – included in formal guidance are subject to diverse and sometimes opposing

interpretations, even within fields. Whether or not such differences of interpretation affect examiners' ability to make sound judgements is unclear, as is the extent to which examiners within a field share a common understanding of the attributes they are seeking in PhD candidates.

The challenges involved in interpreting originality and determining its status in comparison to other attributes were further emphasised by discussions of the focus group, which raised new perspectives. Having confirmed that one of the most important attributes was the candidate's ability to show *'they are able to conduct independent research'* (FG-F3), one contributor suggested that there could be too much emphasis on *'adding something to knowledge or being too original'* (FG-F3), since a candidate could have designed an original study which did not turn out well but they may have *'reported'* it adequately and in such circumstances should be given credit for their novel idea. Another focus group contributor added to this point: *'you learn more by your mistakes, as Einstein said...if things don't work out you're still learning something...either your method was wrong or there is an issue with this problem against your hypothesis...'* (FG-F2).

Candidates' perspectives on originality

A striking finding was that most candidates thought that displaying originality or a contribution to knowledge should be a key attribute of their thesis. When asked what characteristics they thought examiners were seeking in them and their work, most candidates referred explicitly to this key requirement: *'...the only criterion I was really aware of was originality...that you had to make some substantive contribution to knowledge'* (AHSS-C1); *'They make it very clear here from when you start that it has to be an original piece of work that contributes to knowledge'* (STEM-C1); *'So that you could see...that my work was a coherent and original contribution to knowledge'* (AHSS-C5); *'The quality of the work and the originality and how much it adds [to the field]...'* (STEM-C5); *'I was trying to get something that was as original as possible...'* (STEM-C4); *'You should make a contribution...'* (STEM-C3); and *'I would have thought one of [the criteria] is: does this piece of research contribute to our stock of knowledge about that particular area?'* (AHSS-C4).

Candidates' exclusive focus on the need to demonstrate originality and/or a contribution in order to be awarded a PhD contrasted with examiner responses. In most cases the reasons for candidates' certainty are unknown. They may well have been informed by reading institutional

or other guidance, have attended sessions on the doctoral examination organised by their university, or have been guided by their supervisors on the range of attributes sought by examiners. Alternatively their views could have developed from anecdotal evidence about how they would be examined, originality being one of the most familiar attributes associated with the PhD. For example, one candidate with experience of a different discipline *'where people who do PhDs tend to do something completely new [because] there's so much stuff that hasn't been studied'* (AHSS-C1), suggested the research environment had raised their consciousness of originality.

Support available in universities for candidates preparing for the PhD examination appeared variable. One candidate judged their university criteria to be *'very vague'* suggesting they were inadequate in explaining *'what it actually is that they want'* (STEM-C1). This candidate had read other theses completed by graduates in their field to gain insight into what was required to be original and added *'My supervisors would tell me if I'm doing something very wrong'* (STEM-C1), emphasising the key role of supervisors in guiding candidates about the examination.

The data show a range of terms used by examiners to describe the core attribute of successful PhD candidates, including *'originality'*, *'innovation'*, *'novelty'*, *'significance'* or a *'contribution to knowledge'*. Subtle differences can be detected in the ways examiners in the two broad subject groups in this study defined exceptional candidates. For example, all but one of the AHSS examiners implied that there was always some form of *'originality'* in a PhD candidate's work, either because of the methodology, approach or new data combined with their own interpretation, but that a contribution to knowledge that *'shakes the discipline'* was only evident in a few (Mullins and Kiley, 2002). STEM examiners' responses, on the other hand, suggested that they were seeking a contribution to knowledge in every thesis (which AHSS examiners might describe as originality), even if it was *'incremental'* and that true originality or a ground-breaking contribution was rare.

Importantly, respondents agreed that exceptional candidates would display other attributes. This STEM supervisor's response described the differentiation between what most PhD graduates demonstrate and the achievements of an exceptional candidate, linking originality to independent thought, new data and peer reviewed publications:

They should be critical. They should have some independent thought...have some original contribution, and...[candidates should demonstrate] original, independent thinking in the

area and be able to...provide new data...[but] a really good candidate will be somebody who...makes a significant discovery...and [publishes] a very nice paper which would be highly cited. (STEM-S5)

Two questions remain: whether qualifying the 'contribution' in the case of exceptional candidates, using terms such as 'significant' would help to differentiate achievement above the threshold (Kiley, 2009); and if, given that examiners consider originality an essential but not sufficient attribute for the award of a PhD, it would be timely to state this unambiguously. Responses indicate that candidates must simultaneously display a range of quintessential qualities in their work and in person, some of which may evidence exceptional talent and which contribute to their employability. For some, it is the originality of thought and intellectual rigour of the candidate, rather than their ability to produce completely original work, that dominates examiners' judgements. Responses suggest that conflating 'originality' and 'contribution to knowledge' in institutional guidance may be unhelpful because of the variety of interpretations (Clarke and Lunt, 2014), supporting suggestions to abandon 'originality' in favour of variations of 'a contribution to knowledge' (Wellington, 2013; In our time, 2003). The variety of interpretations of originality and a contribution to knowledge, suggest universities have correctly judged that institutional assessment criteria for the PhD can only be meaningful at a higher or more general level. Other research achievements (e.g. the design and methodology, the way in which the research was conducted and situated in the wider field, the analysis, presentation and interpretation of the results), may be best interpreted according to expectations and norms in the relevant field (e.g. the IUBMB's approach). Similarly to originality, the attribute of publishability also represents a range of expectations and is linked with subject, thesis content and structure.

3 Publishability

Examiners and supervisors made frequent references to the 'publishability' of candidates' work, some viewing it as an essential attribute. Interviews and observations suggested that thesis-derived publications had grown in importance, reflecting established priorities in university research cultures (In our time, 2003). The question of whether a candidate's work was publishable might already have been answered prior to the examination, if they had already authored or co-authored a paper in a peer-reviewed journal. The issue of prior publication relates to the multiple thesis models that now exist, with some disciplines expecting work a candidate has previously published to constitute or be included in the thesis. Publishability was interpreted as an indicator of quality (Mullins and Kiley, 2002), allowing

examiners to identify exceptional/differentiate candidates above the PhD examination threshold and to identify exceptional achievement. Prior publications, whether or not included in the thesis, were for some an indicator of the candidate's suitability for entry to the academy.

The significance of publishability is illustrated by frequent direct or indirect references in institutional criteria, which may refer to parts of the thesis meriting publication. Indirect references to publishability include use of the term 'peer review'. Alternatively they may connect the candidate's work with research that would satisfy other scholars in the field, demonstrating the close relationship between the PhD and other peer reviewed publications. Respondents explored the interdependent relationship between peer review and publishability by reference to prior publications and the quality of research in the field.

3.1 Publishability as an indicator of quality and exceptionality

The attribute of publishability, especially its relationship to other characteristics, was employed by examiners and supervisors to differentiate among candidates above the PhD threshold level. Some described contrasting theses, for example this non-case related examiner:

There are some PhDs that you come out and you just say, 'Look, send your dissertation to a publisher'...and there are some where you say 'Well, yeah, okay, it's a pass'...And there's everything in between. (COMP-Ex1)

A supervisor agreed: '*there is a lot of diversity [among] extremely strong theses that you would say...you could publish from, and ones that just about make it...*' (AHSS-S2). Also exploring the relationship between publishability and variability, another examiner described a range in the quality of publications, both in PhD theses and other published work and suggested complementary characteristics:

Broadly speaking...is it publishable work?...Given...we know stuff that's published out there is in a huge range...that's an elastic interpretation, but...what does that mean?...Decent arguments, a story well told, no mistakes...is the work done sufficiently carefully and is it all communicated thus?' (ECON-Ex2)

This confirms that variable quality is found in existing publications, not just among doctoral theses. It also implies that academic peer reviewers of journal articles look for similar characteristics in published work to those sought by PhD examiners, reinforcing the argument that there is a common understanding among academics in a field of knowledge concerning standards, consistency and stewardship of a discipline.

While most candidates did not mention the publishability of their work, its importance as an indicator of quality was well understood by one STEM candidate who, in differentiating between international and 'local' journals, showed awareness of different levels of publication:

The work...has to be publishable otherwise [it] is useless...we should be able to present the work to the world...maybe not to an international journal...[at PhD level] a local journal is also fine. (STEM-C2)

Similarly to examiners, this candidate recognised that expectations of PhD theses might be less exacting than peer review of journal articles.

Discipline-specific influences were reflected in other responses. For example, in a field where candidates may be expected but not required to include previously published work, this examiner suggested the significance of publishability in contributing to academic judgements of thesis quality:

A thesis should be publishable in some format...[it]...should be able to be peer-reviewed by people in our community [who] say, yes, this work is of a good standard, so...if somebody else were to pick up this thesis, they would see at least one or two ...chapters [had been published] in a journal. (STEM-E2)

According to the same examiner, one of several respondents who discussed exceptional candidates, publishability alone did not represent exceptional research achievement. This examiner had already confirmed that in their field, originality was rare, but that they considered all successful candidates had made a contribution to knowledge. However, the examiner did suggest that the presence of both originality and publishability would signify exceptional quality:

Publishable and of good quality doesn't necessarily require it to be novel and original and exciting and changing the world...[it] could be a new method or analysis that is not ...going to hit a high profile journal...equally...a piece of descriptive work, so 'the first time we've shown this or no one's ever looked at this and here is the information about it', is still a valuable scientific contribution...the original critical thinking...and the publishable part...don't have to be linked in any way. If they are, then you're talking about a very high quality thesis and PhD candidate and research scientist because they're able to take what they're doing and then apply it in a big, broader context, and maybe you're going to hit the tabloid journals, which are essentially what every scientist is aiming for...good scientific research that is publishable doesn't necessarily have to be in...high impact journals. (STEM-E2)

This examiner was distinguishing between publishability and a 'valuable scientific contribution', while conveying a complex situation that is particularly relevant to STEM

research and also involves classifying journals according to quality and impact. If a candidate can convey in the thesis the impact of their contribution to knowledge in a wider context, such that it becomes immediately of general interest, this represents a high quality candidate and scientist as well as work that meets the 'contribution to knowledge' criterion.

Others associated publishability with identifying exceptional candidates who would publish from the thesis. This focus group contributor and supervisor from disciplines where candidates are not expected to include previously published work in a thesis suggested:

With dissertations where both examiners are agreed in advance that this is an impressive piece of work, I think the viva can be an exciting opportunity to actually engage with where this might go in publications or in future publications' (FG-M1).

A supervisor alluded to the limited amount of time candidates have to publish while undertaking their PhD: *'Sometimes the whole process of [a PhD] is [to] get it in and make the exceptional contribution through the publications you then produce'* (AHSS-S2).

Combining the inclusion of published work in the thesis and post-examination publication from the thesis, this examiner described a discipline-specific approach that emphasised the critical role of the supervisor:

An exceptional candidate will benefit from the fact that a decent supervisor...will make sure that they get more publications out of the thesis...In a really fast-moving area of research...a supportive supervisor would be helping the candidate to publish the work during the period of the project. (STEM-E5)

PhDs by published work provide a different perspective when compared with publishability as an attribute of PhD candidates and its link with exceptional quality.

3.2 PhDs by published work

One of the cases in this study concerned a PhD by published work in which the candidate relied on artefacts to form the main part of the thesis. The external examiner in this practice-based case contrasted it with theses that include concurrent or recent publications:

This is where publication...is different between the normal PhD and practice [based] PhDs. This [case] was...unusual in two ways: when you...have a student who is much younger, in terms of practice, they've only made one film before probably, then they...make a film which is...not only substantial but is very innovative, which, for the position they're in, is like unbelievable that they've managed to do it and that adds to the debate, discussion, knowledge and the standing of the discipline.' (AHSS-E6)

This examiner argued for the integrity of the PhD by published work, reasoning that the candidate's exceptionalism added to 'the standing of the discipline'. The examiners jointly concluded that this candidate's practice was exceptional. The examiner asserted that in cases such as this the viva was more 'crucial' than in others and that the procedure of submitting by [prior] publication was more challenging

Because the time spent...being registered is very short, the element of proof is really on the publication itself and the thesis is...short and the time spent with the supervisor is short, so the chances of a mishap there are...larger than anywhere else in the PhD system' (AHSS-E6).

While this comment concerned a practice-based study, it is also relevant to other PhDs by published work where a link exists between exceptionalism and publication.

Several examiners and supervisors explored the attribute of publishability from the perspective of the candidate's prior publications, whether or not forming part of the thesis, and their influence on judgements.

3.3 The influence of prior publication on examiners' judgements in STEM subjects

Opinions concerning the influence of prior publication were on a spectrum of expectation, some of the view that they were simply desirable, others acknowledging they increased confidence in the thesis. Not all respondents relied on prior publication to reassure them about the thesis, however, especially in the fields where it remains uncommon for candidates to publish prior to graduation. Data in this section illustrate the influence of prior publications in different contexts and disciplines.

In a field where inclusion of previously published work in the thesis is permitted but not expected, a non case-related examiner agreed it was nevertheless desirable for a candidate to reference their publications:

Probably yes, in the sense that it's not necessarily an expectation but it certainly helps...if... [the candidate has] published some of this work, and...the external can...read the paper and...the thesis and see which bits were [theirs], then that is evidence...[the candidate has] met one of the...key criteria'. (BIO-Ex1)

Another non case-related examiner from a different discipline where the assumption was that a candidate would already have published described how they employed prior publications as an indicator of quality:

I'd [hope]...they would have published at least one peer reviewed journal paper, or a conference paper at the very least [then]...you'd have a good idea that the standard was acceptable within our discipline, assuming it's gone to a journal of appropriate standing...I

always...read the paper...I...use it as an indicator...Often...it's written a bit better in the journal than...in the thesis, which I normally put down to the involvement of other academics...But [publication is] an indicator, I don't rely on it. (ENG-Ex1)

Others confirmed the impact of prior publication in peer-reviewed journals in instilling confidence in candidates' research potential:

If they've had peer-reviewed publication of their work already, that would be a big plus, but I would absolutely expect them to have enough data to make at least one joint first authored publication...that would be the minimum level [although] I have passed someone who hadn't published anything. (STEM-E3).

The same examiner confirmed they would judge the work of someone who had not already published *'with a bit more care'* and again emphasised the role of peer review:

A non peer-reviewed publication is the same as a thesis, but if it's been through a peer review process then you know at least that other outside, non-directly involved people regard this as being original and not fabricated or whatever. I wouldn't say I would dismiss it as being fine and not be critical, but I would be perhaps slightly less alarm bell-y. (STEM-E3)

Two other examiners concurred that prior publication in a peer reviewed journal would put them in *'a better frame of mind'* regarding the thesis. One considered no publications to be problematic; both related publication to standards in the subject:

In many cases, the work of the thesis has already been published in peer-reviewed journals [including in this observed case] so that makes it easier because you...assume that, if their work's been published, it's been peer-reviewed and is...up to a certain standard...If that's not there...that's a major problem. (STEM-I3)

The second concurred: *'If you have published papers before the final PhD assessment, this provides some validation that you have reached the right level'* (STEM-E1). Both emphasised that, for them, prior publication in peer reviewed journals was a highly significant attribute.

A supervisor described the expectation that in their field a candidate would be able to select material from prior publications in peer reviewed journals to form thesis chapters:

Usually...we expect PhD students to have a number of completed studies which could each be a thesis chapter and, alternatively, a publication in a good scientific journal...so...what I expect in my field is a minimum of four...completed experimental studies for which the PhD student has been leading the work and, partially at least, designing it and evaluating it independently. (STEM-S2).

This supervisor linked publication to the candidate's capability for research leadership: having a major role in designing the research and in analysing the data.

This group of examiners and supervisors from fields with a range of expectations concerning prior publication suggested unambiguously that the involvement of independent academic judgement through peer review reassures examiners about the likely quality of the candidate and their work. This was in contrast with AHSS examiners, whose responses may have reflected the convention in many of their subjects to publish from the thesis rather than expecting candidates to include or reference prior publications in it. Not all respondents therefore relied on prior publications for reassurance about the quality of candidates' research, and a few even pinpointed potential disadvantages for candidates in publishing during their PhD.

Most STEM examiners had a positive approach to prior publication, viewing peer approval as an indicator of PhD standard. A few, however, cited time and other pressures as disadvantageous for candidates who published during the PhD. Discipline-specific challenges were at the forefront of the concerns of two external examiners in contrasting STEM fields. The first focused on time challenges for candidates in professional practice:

I wouldn't say [publication] is normal...it's not unknown...quite a few of those I've examined will have a paper [in] the thesis...maybe two...sometimes...even three, but that's unusual...The majority of theses...wouldn't have even a single paper...it's great for the supervisor to be [co-]publishing the work, but...it's much more important for the candidate to get the PhD submitted and then...write the paper. (STEM-E5)

The supervisor in this case suggested only exceptional candidates would have published prior to submission:

Unfortunately not [before submission]...in our system and our experience, it's tended to be the more exceptional candidate who can do that. The majority...struggle to get the thesis done...it's very different to a paper. (STEM-S5).

The external and the supervisor both underscored the challenges inherent in practice-based disciplines, particularly the additional pressures for candidates who attempt to complete the thesis and publish papers while working in professional practice. Suggesting that completing the thesis was more difficult than publishing papers, the supervisor added: *'It's very difficult to convert into a publication...we all want the papers, but [the candidates] need the thesis...'* (STEM-S5)

Another examiner from a different case reported discussing with a colleague two potential disadvantages for candidates with prior publications:

A colleague told me that they find it frustrating when a student has published everything that's in the thesis because there's nothing new to consider, and because it is too late to suggest improvements to publications. (STEM-E1)

This response raised two important considerations. The first challenged the notion of the PhD thesis as a medium for original work when the research ideas and results were already in the public domain. The second illustrated how advice from examiners could be employed to augment the quality of publications emanating from the thesis and was corroborated by the case observations.

A supervisor in another case developed the second argument by concurring that both emerging publications and the thesis can be improved by judicious interventions from examiners:

If a student has already published all the data chapters from their thesis [and it's] been peer-reviewed, am I really going to ask a student to add in a few paragraphs that I think are important? Probably not. Whereas, if I know the student might be writing this up later, I might say 'I think this chapter really needs these things'...If I think they're particularly important, I would say 'I'd really like them in the thesis'...otherwise I might say 'Put them in the paper when you start writing it up'. (STEM-S4)

Another examiner described a feeling of reassurance if the candidate had already published, and explaining why a lack of publications should not disadvantage the candidate:

The impression is substantiated if they have published papers. In this field, it is usual to have got out one or two papers during the PhD...this helps to establish the quality...I'd normally expect the student to have published...but would not judge them if they hadn't...It depends on the type of research the student is doing...in a field-based PhD it would be difficult to write up before the end – there's a lot to do when you get back [whereas]...with desk-based comparative analyses you can send articles to journals while still working on the thesis. (STEM-E1)

This response highlighted constraints for candidates undertaking field work that prevents them from publishing until they have all their data, and emphasised the need for realistic expectations on the part of examiners in such cases.

The evidence in this section, irrespective of subject, confirms that the publishability of a candidate's work, whether during the PhD or post-graduation, has become a critical attribute for those entering academic careers and suggests it has a strong influence on STEM examiners' judgements. Publications may serve to identify 'exceptional' candidates as well as acting as a proxy for the candidate's suitability to be awarded the PhD. Both interview responses and viva observations suggested that examiners commonly offer candidates advice during the viva about publication of their research.

4 Research competence

The PhD signifies that the holder is capable of conducting independent research in their field. This is sometimes explicit in formal assessment criteria and in examiners' shared expectations that the candidate should demonstrate research competence. Indicators included appropriate methodology and research methods, the ability to analyse data, coherence in communicating research results and knowledge of the literature relevant to the candidate's topic. Many of these attributes were identified initially in the thesis, often leading to questions for the candidate in the viva, some of which revealed unexpected information. Attributes concerning research integrity are discussed separately since they encompass a range of personal characteristics wider than research competence. This examiner summarised some of the attributes that suggest research competence:

A...proper...academic way of examining a topic. Is the inquiry substantiated...? [Is] the candidate knowledgeable about the area that they are...adding a piece of work to? Are they arguing...economically and elegantly...Are they using sources properly...[or]...improperly, like using arguments made by someone without citations, so it's a question of plagiarism? You are...also looking...at the range of approaches or methods that they use, so that [the work] is focused but not too narrow...if it's unfocused, that's a weakness, but also, it can't be...so narrow that it doesn't know about something to the left or to the right of their specific argument. (AHSS-E6)

This response exemplified several of the attributes widely considered to demonstrate research competence, showing how the creation of new knowledge is supported by evidence of a range of other attributes.

Research methodology and methods were a priority for some examiners, but were less significant for others. A few respondents did not mention methodology or methods, either because they were considered unimportant, or conversely, so integral to the PhD judgement that they did not need explanation.

4.1 Research methodology and methods

Attributes concerning methodology and research methods vary according to discipline or field of study and are influenced by whether the study is based on generating and analysing qualitative or quantitative data. For example, where qualitative data are concerned, examiners may be more interested in how and why the methodology was chosen and if the research methods enabled the research questions to be answered, whereas in research involving quantitative data, more emphasis may be given to the quality of experimental design including use of appropriate controls, the integrity of the statistical analysis and potential replicability of

the study. However, all examiners were concerned that the methodological approach should be appropriate to the topic, that the research had been conducted with integrity and that the analysis and outcomes could be considered authentic and reliable. To illustrate and further explain the variation, STEM and AHSS approaches to the assessment of the candidate's methodological choices are analysed separately, as follows.

STEM

Examiners in STEM fields tended to focus on technical and practical elements of the research, including any statistical analysis and its relationship to the candidate's findings. This examiner, for example, summarised key criteria for many in STEM fields:

What you're looking for is a combination of technical capability in the subject matter...are the experiments well designed, is the analysis well carried out? Those...[are] core elements. (STEM-I2).

Another took a similar approach, emphasising the importance of the candidate's experimental design in a laboratory-based science:

The quality of the experiment is probably the single most important thing...whether the experiment described makes sense, whether they're controlled properly, that's a key thing, and whether the question that is being asked is somehow sensible. (STEM-I3).

Others made more of the candidate's analysis of their results, particularly use of statistics and how the results had been interpreted, for example:

Whether...the analyses are correct, how they were conducted, if they address the aims, if the candidate has a good understanding of the theoretical background...collected the correct data to address the aims, has done the analysis correctly...interpreted [it] correctly and...in the discussion if [they] show there is a good understanding of what the data actually show, as compared to over-interpreting sets of data. (STEM-I4)

Another STEM examiner referred to taking an interest in the statistics, also touching on originality in the same sentence: *'I probably look at the statistics a little bit – an important way of judging – and then whether it's original, how much work is in there'*. (STEM-E3)

In another case, the viva observation confirmed that the examiners admired the design of the experiments and the integrity of the data, but had concerns about the candidate's knowledge of statistical techniques. This internal examiner in this case confirmed that the

Statistical analysis was inadequate but it wasn't totally clear why that was... was [it] because [the candidate] didn't know how to do it...hadn't thought about it, or made a decision to adopt a particular strategy? (STEM-I2)

As the viva progressed it emerged that the candidate had insufficient knowledge of statistics and that the examiners' feedback would enable the candidate to improve the analysis and presentation of the emerging data. The candidate's supervisor suggested that the examiners had 'picked up' the candidate's 'weakness in defending the statistics...used' (STEM-S2), adding 'It's not so much that [the candidate] used those statistics, but that [they weren't] able to explain why [they] used [those] tests' (STEM-S2). These comments demonstrate the significance of statistical analysis in the presentation of much STEM research. The next examiner's comments also emphasise the need for a good grasp of statistics, while summarising methodological criteria in a particular discipline, including the requirement for replicability:

In...empirical work...methods have to be replicable, so I have to be able to...think...if I...got their data or...went into their forest, I would be able to collect a comparable data set and analyse it the same way...if I can't do that, they need to continue to give me information until I [have] clarity on what their precise research questions and hypotheses are...I like...to trace those...so that in the methods it's clear...how the results are related to those hypotheses...It's nice to see that sort of continuity through an empirical chapter. I'm a bit of a stats Nazi and...will not tolerate pseudo replication. (STEM-E4)

Other STEM examiners used examples from their disciplines to illustrate methodology-related attributes, for example: 'Whether or not they've repeated experiments and so they [have established] that that is the response they get when they put something on cells...is it real?' (BIO-Ex1). This examiner also referred to the importance of controls in experiments and their impact on the integrity of the results, referring to a candidate they had examined:

The controls weren't necessarily...strict, and I wondered whether [the candidate had] actually done the right controls and just not shown them...or whether [they'd]... omitted them...whether or not they understood the importance of the controls...in the experiments. (BIO-Ex1)

The STEM examiners' and supervisors' comments regarding methodology are technical and subject-specific, whereas those of AHSS examiners focus on the importance of methodological choices in successfully addressing research questions irrespective of topic, as explored below.

AHSS

In contrast to terms frequently used by the STEM examiners above (e.g. 'methods', 'technical', 'statistics', 'statistical', 'data', 'experiment'), AHSS examiners quoted in this section (all but one from social sciences fields) tended to refer to 'methodology', 'methodological choices', 'qualitative' or 'quantitative' (social sciences only), or did not mention methodology at all. Two external examiners provided brief statements regarding methodology: 'Is the methodology

appropriate and rigorous? (AHSS-E3) and *'You are obviously also looking...at the range of approaches...that they use, so that it is focused but not too narrow* (AHSS-E6). Neither examiner expanded on their comments to provide more detail about methodology or methods. The methodological concerns of a few other AHSS examiners, on the other hand, were more explicit, one focusing on the multiplicity of choices of methodology and research design open to candidates:

Methodological choices are amongst the most difficult because a candidate chooses [an] ...approach and...they've got to defend it...There obviously would have been a different way of doing it, but...sometimes the candidate is not really sufficiently...on top of...the justification of the methodology. The methodology can be something that is too easily adopted, without proper consideration of alternatives...ultimately...[it must] stand up and...deliver the insights that the candidate is suggesting are there. (AHSS-I1)

In suggesting that making methodological choices in AHSS subjects was challenging because there were alternatives for how a research question might be investigated, this examiner drew attention to a significant difference between subject groups which was further explained by another examiner in social sciences, who gave as an example (unrelated to the cases in this study) a PhD they had examined in anthropology. The examiner suggested that *'anthropological theses don't really feel the need to explain themselves in terms of methods, so there was very little method'* (AHSS-I2) and that this had been a *'real problem'* especially for the co-examiner. The examiners decided that the kind of methodological discussion they had been expecting the candidate's thesis to contain was not appropriate to the disciplinary field in which they were examining. The examiners therefore decided not to raise the issue of methodology in their report since it would not have been justifiable. They acknowledged the candidate's choice of methodology *'was implicit, but you had to read quite carefully to find the account of method'* (AHSS-I2). However, they did discuss the issue with the candidate: *'In the viva, we talked about this, and [the candidate] said that, in [a] previous assessment...the anthropologist told [them they were] worrying too much about the method'* (AHSS-I2).

As the last two examples suggest, in some AHSS fields, choice of methodological options is not always straightforward and may take time to emerge. The anthropology example shows one of the challenges for examiners in interdisciplinary studies where methodological expectations may need to be adjusted to reflect the candidate's field.

In STEM subjects, researchers are potentially constrained by the nature of their research, which might lend itself only to one method, for example, laboratory experiment or structural testing of a design. This is a highly generalised position, however, with many studies defined as

'mixed methods', or involving the generation of both qualitative and quantitative data. Yet as suggested by this examiner from a discipline outside the range of observed cases, qualitative data may provide an important dimension in circumstances where quantitative data are more common:

You are looking for somebody who is using research techniques, which are relevant and best-equipped to answer the questions...If they're asking questions which are value-based...that needs to be reflected in the kind of methodology...they use, and they can't do that...by just doing a quantitative survey. It's got to be something which probes people's understanding...attitudes and backgrounds.' (ECON-Ex1)

The examiner demonstrated the need for research methods tailored to the questions being asked, differentiating between quantitative and qualitative approaches:

In the kinds of question that are not open to...quantitative approaches, the principles remain the same but the...ways in which the evidence is accumulated and discussed and ...the rigour is demonstrated are qualitatively different...[you] need to have a sense from the thesis that here is a person who understands the criteria...to be satisfied in order to demonstrate some notion of truth...the acceptability of the hypothesis is important, at both ends of that methodological spectrum. (ECON-Ex1)

The insight demonstrated in these responses summarised the contrasting approaches in different disciplines while accentuating the need to select a research design and methods that complement the nature of the enquiry. All examiners are concerned with the candidate's choice of research methods and, within broad parameters, judgements are made according to the disciplinary expectations, the candidate's methodological choices and their suitability to investigating research questions.

4.2 Data analysis

The ability to analyse data objectively while making a coherent argument was an attribute sought by STEM examiners and was evident from both interviews and observations. For example, one examiner, having confirmed that '*Quantitative data, quality of analysis, they are...the main things*' (STEM-E4), suggested they would pay particular attention to the rigour of analysis in a case where for whatever reason (e.g. '*if they'd just had an unfortunate time*') a thesis was '*light on data*' (STEM-E4), suggesting the candidate should:

Try and balance it and beef up the analysis side of things so it makes a more substantial piece of work...[the analysis]...should then be excellent, to try and compensate for the fact that they haven't really got that much. (STEM-E4).

This examiner also expected candidates to identify what future research would be needed to build on their own findings. However, they qualified this by suggesting that where data were limited, a possibility not mentioned by AHSS respondents, they might expect the candidate to

have attempted to pursue further research themselves, while acknowledging this might be difficult given time constraints: *'If...it's still a bit light I would...go back and say...actually you need to look at that yourself, rather than just telling people to do it in future'*. In cases where the candidate had adequate data, they could only be expected to suggest areas for further research, since *'It's not realistic to expect you to have done any more within the three-year period'* (STEM-E4).

Another STEM examiner also emphasised the importance of being able to identify further research and the candidate's data analysis abilities:

If they've done the stats correctly...the right statistical analysis and...made the right interpretations...are they actually looking at the data and able to interpret [it]? Are they able to decide where they would go next? Can they identify any key experiments? (BIO-Ex1).

This response illustrates the influence of the topic and field of study on the ways in which examiners' expectations are practically interpreted, the example in this case being laboratory-based science. As these examples show, quantitative data is usually the basis of STEM research, whereas it is possible that the AHSS examiners (mainly in social sciences) did not refer to data analysis as such because they viewed it as part of the attribute of coherence and the ability to make an argument using qualitative data.

4.3 Coherence, including thesis structure and style

The ability to make a coherent argument in the thesis was central to research competence. This examiner, for example, articulated the importance of clarity and coherence: *'I might not necessarily agree with it, but if it's a coherent story, strongly argued, that would be what I would be looking for'* (AHSS-I3).

Other comments were brief, for example: *'...articulate and coherent...'* (EARTH-Ex1), and *'It's helpful if they're lucid and coherent'* (STEM-E5). The next examiner summarised the questions they had in mind that they hoped would be answered from reading the thesis: *'What is [their] main thesis? How have [they] conceptualised the idea? Aspects of the argument'* (AHSS-E3). Another examiner judged these characteristics to be significant, mentioning *'coherence'* (STEM-I3) in particular when referring to the structure and style of the thesis. Finally, an AHSS examiner quoted the institutional criteria provided for doctoral examiners, which included the following: *'...the coherence of the submission, as an academic and a scholarly integrated whole'* (AHSS-I4).

Some referred explicitly to the importance of writing quality, for example: '*Quality of writing*' (AHSS-I2), and describing an integrated approach that included not only thesis presentation but content:

How well written the thesis is, whether it's full of typographical errors from the start; I actually look quite closely at how the figures are presented, how the data is presented... and whether they're properly explained and properly labelled. (STEM-E3)

Examiners' comments indicate that the way in which the thesis is presented, particularly the coherence of the argument, a criterion relating to writing quality, affects their judgement. The last response in particular encapsulates a wide range of both presentational and fundamental attributes that the examiner is expecting candidates to fulfil, such as meticulous attention to detail when presenting their results. This supports the conclusion in Chapter 5 that the thesis is the principal element in forming examiners' judgements, but also illustrates the integrated nature of the assessment process and what the presentation of the thesis tells them about the candidate.

Integration of data was a crucial characteristic for another STEM examiner, who noted the essential role of concluding discussions in drawing strands of the research together to create coherence. The examiner commented on the challenge for candidates in ensuring a logical approach throughout and the need to return to the research questions in the final analysis:

General discussions...can be quite weak because students can...find it hard to bring everything back together and...relate it to those really big questions...It's nice to see...symmetry between the introduction and the discussion, with them...fitting their own work back in to what they discussed in the introduction. (STEM-E4)

The thesis characteristics of integration and coherence are closely linked, especially in making accessible to examiners and others the candidate's contribution to knowledge. As such they are essential components of a thesis and reassure examiners about the candidate's clarity of thought, analytic ability and awareness of the wider field of research.

4.4 Ability to situate research in the wider field

A literature review is a key element of the early stages of training for most doctoral candidates. Data suggest that a comprehensive knowledge of the literature relevant to the candidate's topic is a fundamental expectation of examiners, who seek evidence that candidates can critically appraise their own and others' work, while developing an appreciation of the impact of their own research. Examiners' comments and behaviour suggested that the viva has at least equal weight with the thesis in evaluating candidates' ability to situate their research in the wider field and that this attribute is important irrespective of discipline. This

example from a STEM examiner shows how during the viva a candidate can demonstrate strength in literature knowledge and analysis: *'when I have a PhD student in front of me, I would like somebody who...can make a fool out of me on the literature'* (STEM-I5).

This examiner's comments were representative of others' views in a range of subjects, suggesting how candidates might display their knowledge of the literature:

To be able to answer questions as though they know what they're talking about, as part of the wider field. It's not just about their ability to [analyse] their data and write a thesis...you're looking for them to bring wider knowledge of the field to bear on their research. (STEM-I1)

They confirmed that evaluating these skills was one of the purposes of the viva:

The viva is partly about probing that the candidate has a grasp of the analysis. I'm not sure how I'd respond if they couldn't relate their thesis to the wider body of knowledge, to be able to take a step back from their research and discuss it in a general conversation about the field. (STEM-I1)

This examiner's response articulates a range of criteria that a candidate is expected to fulfil in demonstrating knowledge of the field:

Evidence that the person has...conscientiously and properly mastered the existing literature...you need them to know as much about the...area of the academic literature that they're working on, as anybody else in the world. In some sense...they are the experts...on that body of knowledge, and you would therefore expect an...authoritative and skilful summary of the state of knowledge in their area...it has to be the part of the literature that's germane to what they're doing. (ECON-Ex1)

This respondent was expecting the candidate to have become the international expert in the field pertaining to their topic, but not necessarily the whole subject. Another examiner suggested the candidate should search widely for relevant work in the field: *'To go out and do the relevant literature searches or whatever background reading is necessary, whatever sources, not just literature, to do the background work'* (ENG-Ex1), suggesting that the breadth of the candidate's knowledge *'in surrounding areas not directly linked to their work'* was an important consideration *'to check they've explored the boundaries...or looked at different fields, at similar techniques or whatever'* (ENG-Ex1).

Comments from other STEM examiners confirm that the candidate's ability to situate their own work in the wider body of knowledge is key. One asked if the candidate's work was meaningful in a broader context and confirmed that knowledge of the field was very important. Further STEM examples included: *'I'm looking for...an understanding of the field, or*

the characteristics [of it] – do they understand how the work fits in?’ (STEM-E5); ‘in the discussion...just understanding what the relevance of their results is, what it means that they’ve found and how that fits in with what’s gone before’ (STEM-E4); and ‘Obviously [you want to probe the candidate’s] understanding [of] the impact of the results on the research field, and you usually get that from the discussion’ (STEM-E5).

Some examiners overtly tested the impact of the candidate’s research, for example:

It was a very theoretical PhD so I wanted...to explore how it would be applied in real situations...I deliberately probed...heavily those...areas...How would you use this in a real situation...why is this finding relevant? It was to do with buildings and construction...how can the man on the building site use your findings to better design this building? ...In every case I would tailor it...specifically to the candidate’ (ENG-Ex1).

These responses suggest that successful candidates must convince the examiners of their awareness of the impact of their research and its practical application.

Another examiner linked criteria concerning knowledge of the literature to originality and performance: ‘Ability to...engage in the literature and to present oneself, to...show originality...to show flair’ (AHSS-I2). This theme was developed by another AHSS examiner referencing their role as a supervisor, in differentiating between a descriptive and an analytical approach to the literature:

Not just coverage...but appropriate review of literature...one of the distinguishing features of a good thesis...is an insightful review, rather than a... comprehensive summary of the literature...As a supervisor, I argue that the research questions should emerge almost naturally out of the literature review...once you’ve familiarised the reader with the literature, and highlighted aspects ...of interest...that should provide...the platform for the thesis, where you’re saying this is what others...have done...[these are] the strengths and weaknesses of that research, and these are the questions that have only been partially addressed, therefore...the next step is...what I’m going to do. (AHSS-I1)

In suggesting how the literature review discloses the candidate’s strengths and weaknesses, the examiner explored further the argument that stronger candidates would take an analytical approach:

The literature review can be quite revealing of the strengths or weaknesses of a candidate...the difference...between somebody who simply tells you one thing after another and somebody else who’s actually drawing it together, the beginnings of an argument that you know is going to go on through the evidence...the analysis and the discussion. (AHSS-I1)

In describing the importance of the literature review to introducing an argument that can be sustained throughout the thesis, this examiner made a clear link to analytical ability and

coherence. The differences among candidates as revealed by their approach to the literature was noted by a supervisor in the AHSS group, who similarly linked it to other elements of research competence: *'The ability to handle literature has many grades. There is clear A-grade literature handling...fluency, criticality, argument, those sorts of things'* (AHSS-S1)

Some, such as this examiner, having developed concerns from reading the thesis, take the opportunity to follow up the candidate's thoroughness and knowledge of the literature in the viva, in this case checking that the candidate had read the papers being referenced:

I was...looking at the ability to critically appraise...did [the candidate] actually read those papers [they] cited...then I realised...[the candidate was] citing papers that didn't say what [they] said they said...I happened to know because I looked at the date of the paper and thought 'But they didn't know that [then] – it hadn't been discovered yet'...I wanted to make sure that [the candidate]...had actually read some papers. (BIO-Ex1)

This example also relates to research integrity and the candidate's honesty in citing the literature.

The next examiner's comments underline the role of the theoretical underpinning of the thesis through the literature review in demonstrating the candidate's understanding of the wider field and where their research is situated using subject-specific references:

I'm looking for a good theoretical understanding...how their empirical work is going to contribute to the theory...There's going to be massive variation in how much theory there is, depending on the topic – if it's a welfare-related one there's...virtually no theory, but in [this case], there's a...heap of communication theory and evolution of language theory that [they] did a very good job of reviewing...showing [they] really understood it...A solid understanding of the most relevant previous empirical research...what I really like to see is that they use that and it's not just...descriptive...but that they...build a really strong case and a rationale to say why their thesis is important and...fits in with what has gone previously (STEM-E4).

This and other respondents underlined the importance of the candidate taking an analytical approach to existing theory in building the argument for their research.

4.5 Research integrity

The significance of ethical research practice was evident from responses in this study. Several examiners touched on criteria concerning ethical behaviour and candidate integrity, including avoidance of plagiarism. Some also raised the question of supervisor input to the thesis and the need for candidates to acknowledge other authors, with the extent of the supervisor's input to the thesis treated separately from how to represent multiple authorship of published papers presented as thesis chapters. The expectation is that the qualities of integrity and

honesty should be identifiable in all doctoral graduates and post-doctoral researchers (EUA-CDE 2016).

Several examiners briefly referenced characteristics demonstrating the candidate's personal integrity, for example: *'This might seem like a bit of a strange one, but I like to think that they're honest'* (STEM-E5); *'Thoughtfulness, honesty, modesty'* (AHSS-I2); and *'...and honesty.'* (STEM-I3). A non case-related examiner provided an interpretation of honesty through relating it directly to the creation and exploration of research data:

One would not want to be judging a thesis where the person had been allowed to ask trivial questions...particularly in qualitative work, it is possible to ask questions that...are not terribly interesting or important...or new...you're looking really for...honesty to say...this is what we know, and this is actually an interesting way to develop [it]...you're looking for...honest commitment about the research questions that are being posed (ECON-Ex1).

The principal attribute concerning research integrity was the avoidance of plagiarism, but in all cases examiners sought the candidate's personal integrity in conducting research, their awareness of the need to acknowledge others' input and the extent of their contribution to the research and the thesis. Examiners were also evaluating the candidate's ability to work independently and lead the research process rather than being led by their supervisors.

Avoiding plagiarism

Learning how to cite other authors appropriately is a key element of research training. As well as being one of the key purposes of the viva, testing the candidate's ownership of the work plays a part in checking their personal integrity, as suggested by the first two examiners: *'Well, first of all, I want to know if they've done the work'* (STEM-E5); and:

Have they done the work themselves? Have they acknowledged where other people's work had an input, because it's bound to, so their honesty and their integrity with respect to what they're presenting. (BIO-Ex1).

Two examiners linked ownership of the work to exploring the data analysis with candidates during the viva:

They...need to demonstrate the work is their own when discussing the analysis of data – how they have done it. Once I asked a student about the statistical analysis...they said someone else had done it and they hadn't understood what the person had done. (STEM-I1)

The second also revealed that in one viva (unobserved) they had discovered that the candidate had not undertaken their own statistical analysis:

Do they sound as though they did the work is the very basic one...[or]...as though they've been heavily reliant on other people and so are a bit at sea in the viva situation? To give an extreme example of that, I had someone who said 'Well, someone else worked out all the statistics for me so, if you're going to ask me questions, I won't necessarily be able to answer [those] on the statistics'. (AHSS-E1)

Two more examiners, on the other hand, were inclined to take as read the candidate's honesty and integrity: '*I just assume that everyone's work is their own. I'd be horrified if I went to a viva and that wasn't so*' (STEM E4); and '*There is an element of trust there that they've done what they've said they've done...*' (AHSS-I3). These positions were in the minority, especially when compared with those who saw the viva as an important opportunity to authenticate the thesis.

A central feature of avoiding plagiarism was the acknowledgement of any input from other authors, which was uncontroversial. As the next section shows, however, there is no consensus concerning the representation of supervisor input to the thesis.

Acknowledging contributions from supervisors and other authors

Opinions were divided on how and how much a supervisor could and should help a candidate to write the thesis and to some extent they reflected collaborative and non-collaborative research practice in different fields. Positions ranged from acceptance that in a particular field research was now a collaborative venture (Morley, 2004), to a strict policy that the thesis should only be the student's work.

A STEM examiner, commenting on the nature of collaborative research in their field, leading to multiple authorship of publications, suggested that co-writing with the supervisor was inevitable and that the viva enabled examiners to estimate the extent of the candidate's contribution:

My view is that the work will be co-written with the supervisor. I don't expect anything else...that's implicit, because in the end that's going to be published...it won't be [the candidate] alone, it's [them] with the supervisor and some others...most science today is collaborative...you're working with other people...you're learning from them...When you see a thesis - who actually wrote it? Who thought of this? ...We shouldn't be worried about it, but that's why having a couple of hours with the person alone is a good idea. No supervisor, just the candidate. (STEM-I2)

These comments re-emphasise the importance of the viva in determining authorship and reflect the critical role of the supervisor in the candidate's doctoral training. They are particularly relevant to the convention in some science subjects for candidates to submit as part of their thesis co-authored papers relating to the study. Examiner STEM-I2 raised the

question of how far the ideas presented in the thesis were the candidate's own or those of their supervisors and how the balance of authorship is managed in published papers, including any included as a thesis chapter. It could be argued that the examiner is presenting a pragmatic view of how authorship is assigned to research output, recognising that the ideas of several individuals (including co-authors and supervisors) influence and improve the final outcome. Other STEM examiners supported this approach, one raising the question of team-based research when discussing the purposes of the viva: *'Do I get a sense that the candidate has been working alone or as part of a team, and I can obviously establish that later at the viva'* (STEM-E5). Another examiner acknowledged the collaborative nature of research and, similarly to one of the candidates (STEM-C3), emphasised the importance of the candidate leading the research:

A lot of research is collaborative so supervisors will have had input into how [the student's] research has developed. If there were ever any suggestion the student had just been a research assistant for the supervisor, I wouldn't accept it. (STEM-E1)

Another agreed there was a need to test the extent of the candidate's contribution:

You have to make sure that the thesis has really been written by the candidate and is not just written by the supervisor, and that the work's all been done by the candidate. (STEM-I3)

On the other hand, some examiners and supervisors were less sanguine about the extent of the supervisor's contribution to writing the thesis, as in the following three examples, the first from a supervisor drawing on their examining experience:

You want to see...evidence that...students can explain what they've done, why they've done it, how they've done it, how they interpret data and so on...you don't necessarily get that from reading the thesis, but you do get it from talking to students...I've had cases where...as an...examiner...I've pointed at a figure...asked the student to explain [it]...and they've said, 'I've no idea – my supervisor gave me that figure'...that's not what you want. (STEM-S2)

This supervisor illustrated how both elements of the examination play different roles in the process and how, without a viva, it could be possible for a supervisor's input to a thesis to be confused with the candidate's own knowledge and understanding.

Two STEM external examiners confirmed the extent of supervisor input was something they tested in the viva:

There might be a group...you're not convinced they're going to be an independent PI type person, but who are competent, they've executed good work, and it's been done to a high standard. You want to get a feeling for how much was their own work and how

much was their supervisor's...It doesn't necessarily determine whether they pass or fail but it might determine your overall view of them. (STEM-E3)

I...worry...about where...supervisors get too involved and actually land up editing theses more than they should...that's why...the viva is important, to ensure that the student understands what's written in the thesis and...has actually written that themselves. (STEM-E5)

An AHSS supervisor spoke about the thesis of an international candidate they had supervised whose competence in written English was poor. The supervisor was satisfied with the intellectual achievement of the candidate, so spent much time and effort correcting the English in the thesis. The supervisor reported how:

The examiner did actually express great surprise that this beautifully-written text was produced by that...very limited English student who he met in the viva... so worried was he that he [asked] me about it. (AHSS-S1)

This supervisor also differentiated between forms of help with the thesis:

One is the English language thing, but the second is a sort of...academic thing, and I suppose that there are some supervisors who make quite a contribution to the writing of the thesis. (AHSS-S1)

Since one of the criteria employed by examiners in assessing PhD candidates is the ability to conduct independent research, the student's ability to lead a research project and make difficult decisions if the research does not run smoothly is critical to the judgements made in the examination process as well as to the individual's professional training. One external examiner confirmed that evidence of the student's independent thought was important during the assessment process: *'I'm also looking for the student's research to be independent – their own, rather than their supervisors' – using ideas they've largely developed themselves'* (STEM-E1).

The final comment below connects the candidate's personal integrity with the ability to work independently and take the initiative in their research, thereby integrating both elements of this section:

At the start of the viva, we had a little discrepancy of whether [the candidate had] contributed all or part or some of the work...I wanted to see evidence that [they] understood all of that and that it was really [the candidate] driving [the research]. (STEM-E2)

In this case, the co-author(s) of one of the thesis chapters had not been acknowledged (the chapter had already been published as a journal paper and was in the public domain so there was no doubt about authorship), but the examiners were clear in their feedback that this was

a mistake, that co-authors needed to be cited, and that awareness of this convention was part of the candidate's training to become an independent researcher.

The attributes explored in this section show different facets of research competence, some of which are more important than others to examiners in AHSS or STEM fields, with others appearing to have universal resonance (e.g. ability to situate research in the wider field). Attributes integral to research competence are closely related to research integrity, a significant and increasingly important characteristic of researchers (EUA-CDE, 2016). Data concerning these and other elements of ethical research suggest the complexity of evaluating the many facets of research integrity, a personal characteristic of doctoral candidates that on graduation has potential to extend to other contexts. The extent to which this and other personal characteristics have become substantive in examiners' judgements is explored in the next section on intellectual rigour, another multi-faceted attribute that is often a priority for examiners.

5 Intellectual rigour

In the context of the PhD examination 'intellectual rigour' has resonance across subjects and contexts. It is an all-encompassing term representing qualities that demonstrate thoroughness and independence in the approach to research. As a preamble to exploring detailed attributes, some examiners used the expression to represent, for example, the quality of argument evident in the thesis and the candidate's ability to respond satisfactorily to challenging questions in the viva. An intellectually rigorous approach was also linked to the candidate's capacity to think quickly when asked unexpected questions in the viva: to 'think on their feet' and to display agility in demonstrating depth and breadth of knowledge of their own research and the wider field. The concept of intellectual rigour extends to the capacity for independent thought, and the ability to lead a complex research project. Formal guidance frequently includes the capability to conduct independent research as one of the principal criteria for doctoral assessment. The examiners' comments in this section reference the thesis and viva: some specifically address one or other element, while others do not differentiate, suggesting that both elements are important in assessing the characteristics in question.

One examiner gave a concise definition: *'A whole set of things...go together...they ultimately come down to what might be called intellectual rigour, in that the evidence and the argument stands up to challenge, and within that, there are a whole number of different attributes'*

(AHSS-I1). Another referred to rigorous thinking in the context of maintaining discipline-specific standards:

My...expectations [are that]...the way they're thinking is rigorous. They don't have to think the same way as I do, but they do have to do it to my standards...the quality of their thinking and the analysis etc. (EARTH-Ex1),

Another non-case related examiner in a different field, suggested: *'Two things are particularly important...the most important really is rigour'* (ECON-Ex1). The second 'important' factor for this examiner was that the candidate should demonstrate originality, but this was secondary to their first priority. In this section, a range of qualities cited by examiners suggest agreement about key criteria while also highlighting differences in terminology.

5.1 Critical, analytical and reflective thinking

The capacity for critical, analytical thinking, together with a logical and reflective approach to research, were characteristics cited by several examiners in relation to intellectual rigour. Attributes such as these were described in terms of their relevance to the candidate's research achievements. However, they also have wider value in relation to the candidate's plans for the future irrespective of environment. For example, a willingness to be self-critical and to reflect on one's own practice are potentially useful qualities in many contexts.

Examiners' expectations were that candidates should be able to apply their critical skills to their own and others' work, for example: *'They should be critical...'* (STEM-S5); *'Willingness to take a position...self-critical and critical, in the best possible academic sense'* (AHSS-I2) and:

If they're doing empirical work...the ability to be reflective is really important, because obviously, we don't see them in their interactions with participants, we see their final written-up piece. (AHSS-I3)

The third examiner seems to be suggesting that the viva is their opportunity to evaluate the candidate's ability for reflection, since it is not possible to gauge this from the written work.

Others stressed the importance of using theory in the analysis of the research, for example this internal examiner when discussing the viva: *'The candidate's ability to think analytically, to bring in theory in responding to the examiners' questions, to bring theory and analysis together'* (STEM-I1). Another combined the need for the candidate to be self-critical and constructively critical of others' work when analysing their own results:

Are they questioning...challenging what they themselves have done, and what others have done [by] saying...this is what I've found and therefore I disagree with the work that's done by others...they need to show they're analytical. (STEM-E5)

Having discussed the creation of new knowledge, another examiner referenced self-criticality regarding the candidate's understanding the limitations of their research, implying that this quality could best be evaluated in the viva:

The next...thing...is how self-critical the candidate can be...do they understand the limitations? Often the thesis is written in a very positive way, which is not surprising – you want to put your story across, that's how you publish. But we want to know [if] you realise that...there are always things that...you may not have proved sufficiently, [so] what will you do more to prove it. (STEM-I3)

The need for candidates to be aware of imperfections in their own work was also cited by this examiner:

Do they acknowledge any deficiencies in their work...are they able to say 'In retrospect I would have done this [or] that' because it is after all a learning process, so you're not looking for perfection at the outset. (BIO-Ex1)

Describing undertaking a PhD as a learning process is especially relevant to its often-cited purpose as an award signifying capability of undertaking research independently. Defining the kind of critical thinking expected of PhD candidates in general, the next examiner introduced the concept of deep and surface learning more commonly discussed in the context of taught degrees, yet also appropriate here:

I look for...deeper critical thinking and understanding of the subject, so that it's not just surface learning of the topic and...understanding...the basics. It's being able to take those...integrate [them] and try and...[show]...that they are an expert in the topic...At the end, you should be a researcher who is...aiming towards being an expert in that topic...holistic understanding and...critical thinking about the work you're doing is [the] real one...and I have that in mind definitely during the viva. (STEM-E2)

This examiner articulates particularly clearly the expectation that becoming an 'expert' in the topic signifies the candidate has demonstrated intellectual rigour in their grasp of the wider field in which they situate their own research, confirming that evaluating the composite qualities that demonstrate such rigour is uppermost during the viva.

The characteristics of critical, analytical and reflective thinking were mostly but not exclusively referenced by STEM examiners. They suggest that candidates' capacity for analysis and to be constructively critical of their own and others' work is integral to the PhD judgement.

5.2 Problem-solving and logic

Other characteristics exemplifying intellectual rigour include problem solving and disciplined thinking. For this examiner, for example, intellectual rigour was represented by:

An enquiring mind; to be able to evaluate exactly what the problem is that needs to be solved...establish that there is a need for further research; and be able to demonstrate that they can do that research themselves to a sufficient standard. (ENG-Ex1)

Another STEM examiner specifically referred to the need for logic: *'I like to see that they're logical, that they can work their way through a problem...they're analysing and...do that in a logical manner'* (STEM-E5).

A supervisor social sciences also referred explicitly to problem-solving ability but as part of a set of attributes that would differentiate an excellent candidate from 'middle of the road':

In the excellent candidate...you're looking for...research agility and flexibility and ability to problem-solve, to think on their feet, to think about new problems, new solutions...think divergently, creatively. (AHSS-S1)

Other examiners also saw the viva as an opportunity for testing the candidate's ability to respond well to questions they had not necessarily prepared for. For example: *'really importantly...are they capable of thinking on their feet?'* (STEM-E5); and:

The advantage of...the viva is that you're asking them questions on the hoof. However much they've been coached or prepared they wouldn't be able to do that unless they're competent. (STEM-E3)

Another asked *'Do they have the capacity at the viva to engage in intellectual defence of the thesis – prepared to respond to questions they haven't seen in advance?'* (AHSS-E3). All three examiners suggest that the ability to respond effectively to questions in the viva, including detailed enquiries about data analysis, is a measure of the candidate's intellect.

STEM examiners in particular sought evidence of a candidate's independent thought and ability to lead research rather than follow others.

5.3 Independent thought and research leadership

Institutional and other guidance enshrined in doctoral assessment criteria frequently includes reference to the 'ability to conduct independent research' as one of the key characteristics of doctoral graduates (Table 2.1). Some examiners and supervisors further defined this abstract concept to provide insight into how it is practically interpreted in the assessment process:

The extent to which someone is in the driver's seat for the research that they've done during their PhD...ultimately, whether someone succeeds as a scientist is [about]...can they ask their own questions, design their own ways to answer...questions, and then take charge of...converging all...this into a final product by way of a scientific publication? ...That...isn't necessarily apparent from...looking at a thesis because as an outsider,

you...don't know...the extent to which a student has designed the work, evaluated the data by themselves, and indeed written it up. (STEM-S2)

The ability of the candidate to take the initiative in their research is also significant for this examiner who touched on a characteristic exhibited by several candidates in this study – the expectation that they will have ideas for the direction of future research into their topic:

[To] bring out the fact that they have an enquiring research mind so that they're capable of formulating their own theories...of going off and doing relevant background investigations and formulating a proper plan for attacking the problem, for taking it further. (ENG-Ex1)

These two responses conveyed the importance of independent thought and leadership of research by suggesting a range of characteristics a successful candidate would display. Others were less descriptive, but referred to independence: '*They should have some independent thought*' (STEM-S5); and '*It's very important...the candidate shows that they are able to conduct independent research*' (FG-F3).

Another examiner reflected on the extent of the candidate's independence to make a comparison between those with the qualities required to become a successful research scientist and those who simply reach the required standard:

At the top end you're looking to say 'Is this person going to make a competent and capable, independent...scientist?'...Would you be happy for this person to run a lab...do they have the...intellect, the rigour...do they have the integrity?' But there's a...let-out clause: this person may not be an independent scientist but have they completed a training period? Do they have enough skills to warrant the award of a title after their name? (STEM-E3)

The views expressed here suggest a range of achievement and expectations above the PhD threshold and that several of the qualities which represent intellectual rigour 'at the top end' including independent thought, would be valuable in any career involving leadership and requiring high levels of integrity.

Responses concerning intellectual rigour emphasise examiners' expectations that successful candidates possess a range of attributes that, while evaluated through the prism of the candidate's research output, illustrate personal characteristics such as logic, integrity, the ability to think on their feet and be reflective, that are transferable to other contexts. Some suggested these attributes were evaluated in the viva. Most did not overtly suggest that intellectual rigour was a valuable attribute more generally, but the characteristics they cited could be described as useful life skills that are potentially beneficial in a wide range of careers as well as in society and personally, as recognised by this examiner:

Are they going to go on to be a successful scientist? Would I recommend them to someone else if they wanted a post-doctoral position? But...maybe they don't want a scientific career, and that's perfectly OK as well...There, you're looking for...are they competent for other things in life? Are they going to bring the skills that they've learned from doing this doctoral research?...It's not just an academic business; quite a lot, well, the majority...are not going on to do post-doctoral work in universities. (STEM-I2)

One of the most interesting features of this quotation is that, although throughout their interview and during the viva I observed, the examiner focused on the candidate as a scientist and the attributes they would need in pursuing a scientific career, here they confirmed that they had in mind the attributes were transferable to other contexts, including employment.

6 Conclusion

The range and significance of the attributes explored in this chapter suggest that three important conclusions can be drawn. First, it is clear that originality or a contribution to knowledge is necessary but not sufficient for the award of a PhD; second, while some attributes are research-specific and others professional and personal, they may be so amalgamated in the individual that it is difficult to separate them; and third, both thesis and viva are significant in examiners' evolving judgements. Chapter 5 shows that examiners give prominence to the thesis when forming judgements. It was therefore striking that many responses implied such a close relationship between the thesis and the viva in identifying and evaluating candidate attributes. For example, some of the sought-after attributes contributing to intellectual rigour, while judged through research achievements, concern both the candidate's work and their professional and personal characteristics and are assessed throughout, but particularly in the viva.

Data also suggest some consistency among examiners in the attributes sought. A range of essential attributes, including intellectual rigour and research integrity exemplified for many a mix of research-specific and personal qualities that would be evident in successful candidates (or those who possess 'doctorateness?'). Evidence concerning examiners' expectations of intellectual rigour resonates with Mullins and Kiley's (2002) findings, whose respondents also cited qualities such as originality, coherence, logical argument and independence in candidates who passed. Interpretation of the attribute of originality, however, was challenging and complex for examiners. This and publishability were characterised by contradictory approaches both between and within fields, though there was some agreement about indicators of exceptional candidates who some identified by the presence of both originality

and publishability. The emphasis on publishability by STEM respondents was striking, preferences demonstrating differences between fields. While unsurprising, it was useful to have evidence of differences between STEM and AHSS respondents concerning research methodology.

Omissions can be as significant as what is included: few of my respondents referred explicitly to what could be described as transferable skills when describing candidate attributes, whereas Nyquist's (2002) taxonomy of 'Core competencies of successful PhDs' focus equally on research and transferable skills. Table 6.1 compares attributes identified in this study and Nyquist's 'core competencies' (ibid) that emerged from the 'Re-envisioning the PhD' study undertaken in the United States (Nyquist and Woodford, 2000) and was based on contributions from a sample of individuals 'who are obtaining, supervising, funding and employing PhDs' (ibid). While Nyquist and Woodford's study is based on suggestions for 're-envisioning' the PhD programme and 'the characteristics of those who are currently successfully navigating a variety of careers after receiving a PhD' (Nyquist, 2002), my contributors (and respondents who contributed to the Clarke and Lunt paper) were asked about 'core' attributes of PhD graduates at the point of graduation. As Table 6.1 shows, though Nyquist's 'generic set of core competencies for doctoral recipients', partly align with my data, the competencies include more wide-ranging attributes relevant to employment but do not include publishability:

This study: candidate attributes	Nyquist, 2002 'core competencies'
Originality, innovation, novelty, significance or a contribution to knowledge	<i>Disciplinary knowledge – what is known, plus creative and adventurous ways of discovering new knowledge, the foundation of the PhD</i>
Publishability: <ul style="list-style-type: none"> As an indicator of quality, of exceptionality and as demonstrated through PhD by published work. 	N/A
Research competence and integrity: <ul style="list-style-type: none"> Research methodology and methods Data analysis Coherence, including thesis structure Ability to situate research in the wider field Avoiding plagiarism and acknowledging contributions from others 	<i>Ability and preparedness to connect one's work to that of others, within and across disciplines...</i> <i>Understanding of ethical conduct as researchers...including issues of intellectual property</i>

Intellectual rigour: <ul style="list-style-type: none"> • Critical, analytical and reflective thinking • Problem-solving and logic • Independent thought and research leadership 	<i>...preparedness to be a leader... Understanding of the mentoring process necessary to provide leadership for future generations in academic or the workplace</i>
Additional 'core competencies' identified by Nyquist but not prioritised by contributors to this study	
<i>Commitment to an informed career choice based on exposure to a broad array of opportunities and paths</i>	
<i>Teaching competency, broadly considered – in one-to-one interactions in the classroom...</i>	
<i>...preparedness to be ...a faculty member, a project manager, a motivator and an evaluator of others' learning in the government, non-profit, corporate or academic sectors</i>	
<i>Understanding of the diversity of present and future students and present and future workforces</i>	
<i>Global perspective – the importance of doctoral work in relation to a global economy, sensitivity to cultural differences</i>	
<i>Ability to see oneself as a scholar-citizen who will connect his or her expertise to the needs of society</i>	
<i>Ability to communicate and work in teams and explain work to public audiences and to those who set policies</i>	

Table 6.1: Comparison of attributes identified in this study with Nyquist's 'core competencies' (2002)

It was to be expected that some of Nyquist's 'stakeholders' who included government agencies, non-profit organisations, business and industry, as well as academic organisations, would place greater emphasis on competence in the wider context, while in the other studies contributors from the academic environment focused mainly on research-related qualities at the point of graduation, which was evident from observations. However, it is interesting that Nyquist's contributors have subtly linked attributes often developed through research to common demands in the workplace, whether academic or other, while these are rarely mentioned by my respondents. The 'teaching competency' category is less surprising in a North American study, where all PhDs are involved in teaching.

The four groups of attributes considered by my examiner respondents as central to their judgements suggests they share much common ground. Table 6.2 presents these attributes, suggesting if they are apparent from the thesis, viva, or both and whether they display research or professional skills, or both. The primary source of evidence is in bold. The most striking feature of this analysis is the extent to which the role of each element complements the other. For example, while the thesis contributes most to evaluating the research-centred attributes of originality, publishability and research competence, the viva also plays a part. Conversely, the viva makes a more significant contribution to assessing the core professional and personal attributes contributing to the intellectual rigour attributes, although some of

these are also demonstrated in the thesis. Table 6.2 shows that apart from A2, both elements contribute to assessing each group of attributes.

<i>Attribute group and sub-groups</i>	<i>Apparent from thesis (T) or viva (V)</i>	<i>Research (R) or professional/personal (P) skills</i>
A1 Originality / contribution to knowledge Essential attribute. Candidates were aware of its importance. Different between STEM and AHSS: in STEM most candidates make `contribution` but few are `original`; in AHSS most display `originality` but only exceptional candidates make a `contribution`.	T V	R P (intellectual rigour, independent thought)
A2 Publishability For some, equally as important as originality. AHSS want work to be publishable but fewer expect candidates to have published during PhD; in STEM, some expect publications in thesis.	T	R P (communication, make an argument)
A3 Research competence STEM examiners focused on technical capability: experiment design, data interpretation, reproducibility, analysis of results; AHSS examiners were more concerned with choice of methodology and if methods had enabled them to answer research question. All sought research integrity.	T V	R P (creativity, flair, initiative, breadth)
A4 Intellectual rigour Attributes included: - Critical, analytical and reflective thinking - Problem-solving and logic - Independent thought and research leadership No subject variations; integration between research and professional attributes was such that they were often indistinguishable.	V T	P R

Table 6.2: Attributes sought and where they are identified

The attributes align well with Lovitts' `technical` (research) and `indeterminate` (professional and personal) qualities, but my findings go further in demonstrating how, similarly to the thesis and viva, the attributes are inextricably linked in examiners' judgements.

This chapter confirms that assessment of the PhD is a multi-layered and complex process, where examiners come to a holistic judgement about the candidate and their work, frequently blurring the boundaries between research and professional and personal characteristics. The range of attributes therefore includes characteristics relevant to both academic and non-academic contexts, confirming the thesis and the viva play respectively fundamental roles in assessing candidate attributes. Evidence suggests that to be judged able to conduct research independently, a candidate must demonstrate a high level of research achievement or technical ability (Delamont et al., 2002; Lovitts, 2007), together with a range of professional

and personal characteristics. Acquisition of a wide range of attributes through the process of doctoral training benefits doctoral graduates in other contexts, irrespective of career path, as exemplified by one STEM examiner, who suggested that *'doing the PhD gives you'* skills that are valued in *'any sort of corporate world'* and that employers *'are after people who can organise themselves, organise other people, organise their work in such a way that it makes sense and...can present their work to people in oral, written, whatever format'* (STEM-E2). The examiner concluded that at the end of the PhD *'those skills have just skyrocketed'*. These are outcomes that require greater accentuation in the context of employability.

Chapter 7: Examination outcomes: diversity versus comparability, and grading

We need...to recognise assessment as a social product, in which the values and traditions of particular cultures and the interests of...groups within them combine to produce particular definitions of quality or merit. (Broadfoot, 2000: xii)

1 Introduction

The distinctiveness of the PhD examination compared with other assessments in higher education is demonstrated as much by its outcomes as by the process itself and the complex array of candidate attributes sought by examiners. Characteristics such as the individuality, privacy and lack of anonymity of the two-part examination contribute to scepticism about the consistency and reliability of examiners' judgements. Additionally and possibly of equal significance, the freedom of universities to determine their own range of outcomes as illustrated in Table 7.1 exacerbates the potential variation in standards already inherent in the range of options available to examiners, especially when coupled with the variable use of time periods within which corrections must be completed.

Potential outcome	Time period (F/T)	Comments
Pass with no corrections/amendments	N/A	N/A
Pass subject to completion of corrections /amendments	Up to 4 weeks	Only offered by few institutions, for the correction of presentational or typographical errors only.
Pass subject to completion of minor corrections /amendments	4 – 12 weeks The exact period may be open to the examiners' discretion although may require permission of the university administration	Much variation within this category that includes a range of time periods. Most institutions offer extensions to initial deadlines for completing minor corrections of typographical errors or diagrams. Some also allow more significant corrections if they do not affect the intellectual content.
Pass subject to major corrections /amendments	6 – 9 months	In this category, no re-examination is required, just resubmission of the thesis with amendments approved by all examiners.
Referral /Resubmission	12 – 18 months	This category is normally separated from major corrections. The candidate is required to substantively revise the thesis with significant guidance from examiners and supervisors and re-submit it for examination either with or without a viva. All examiners must be satisfied with the final thesis.

Award of lower degree	N/A	Normally a research masters degree, e.g. MPhil, this may be awarded either immediately after the first assessment or as a result of resubmission and a second viva.
Fail	N/A	N/A
<p>Time periods: institutions may make provision for different time periods depending on whether the candidate is full or part time, usually allowing part-time candidates up to twice as long to complete corrections.</p> <p>Lower awards and failure: any candidate who wishes to may appeal against the recommendation for the award of a lower degree, or a fail outcome.</p>		

Table 7.1: Summary of common PhD examination outcomes and time periods within which corrections are completed

While the categories for recommendations may be broadly similar, for example ‘minor’ and ‘major’ corrections or amendments are terms used by many institutions, they are not standardised. Appendix 6, section 3, contains examples of university regulations that specify the outcomes available to examiners and procedures to be followed at the end of the examination.

The pass/fail summative judgement in the final doctoral examination masks the diversity of attainment among PhD graduates that is initially reflected in the conditional outcomes available to examiners. Options include failing the candidate or awarding a lower degree, both of which are rare. Occasionally candidates pass with no corrections. Other options include ‘referral’, often linked to major corrections, and sometimes including re-submission of the thesis, either with (rarely) or without another viva taking place. Arguably the range of options available increases the complexity of the judgement yet enables examiners to reflect candidate individuality in their recommendations, allowing for the uncertain outcomes of research and the researcher’s achievements, attributes and circumstances to be reflected in the preliminary judgement.

Once any corrections have been completed to the examiners’ satisfaction and ratified by the relevant university committee, candidates are judged to have reached the required ‘standard’ to be awarded the PhD. The question: ‘What is a high enough standard and is it similar in every subject?’ presents a long-standing challenge for universities and policy-makers. Most universities provide guidance for PhD examiners on the high-level attributes examiners should assure themselves a successful candidate possesses, such as those in the PhD assessment regulations at Appendix 6. For example: ‘*the student has made a significant and substantial contribution in the particular field of learning within which the subject of the thesis falls*’. The

Oxford University regulations include an additional statement that in part addresses the question of whether a candidate has 'done enough' to be awarded the degree, as follows:

Examiners shall bear in mind that their judgement of the substantial significance of the work should take into account what may reasonably be expected of a capable and diligent student after three or at most four years of full-time study in the case of a full-time student, or eight years in the case of a part-time student. (Oxford University)

The inclusion of this advice in the regulations suggests that, as long as their expectations are reasonable, examiners in the relevant subject will be able to determine a 'threshold' level that candidates should attain within the period allowed. This encourages examiners to be fair while maintaining standards in the PhD, yet its relative vagueness encapsulates the difficulties inherent in attempting to define a doctoral 'level'. Nor does it explicitly reflect the range of attainment evident from the data in this study that suggests various levels above the 'threshold' and a low failure rate, the latter partially explained by the progression hurdles now in place for most UK PhD candidates. Given the challenges of defining 'doctorateness' and the high stakes summative judgement examiners are making in the PhD assessment, the reliability and integrity of examination outcomes is central to questions concerning consistency of standards.

Evidence in this chapter is derived from 38 (88%) of my 43 interviewees and focus group members as follows: nine external and eight internal examiners; eight candidates; five supervisors, five non case-related examiners one convenor and two focus group members. I also referred to regulations concerning examination outcomes at several universities.

Data suggest that a considerable amount of diversity exists among those who pass the examination, but that the threshold judgement nevertheless reflects some comparability of graduate attainment. The research question on which this chapter is based: '*How do different examination outcomes reflect candidates' achievements, and is a pass/fail judgement still appropriate?*' emerged from the lack of differentiation above the threshold. In the first section I explore the range of attainment identified in doctoral candidates and the diversity and comparability among those who pass, from outstanding achievement to near failure. In the second, I present data concerning the approximate number of candidates asked to complete major corrections, awarded a lower degree, or failing, as experienced by respondents. Numerical summaries are supported by a range of views on the different examination outcomes, for example: why failure occurs; if fewer or more candidates should fail or be awarded a lower degree; and the relationship between candidate outcomes and employment

aspirations. The final section addresses the benefits and disadvantages of grading the UK PhD, including the possibility of a distinction to reward exceptional performance.

2 The range of attainment

The wide range of attainment evident from this study is only partially reflected in the recommendations available to examiners. This variation arguably allows for the individuality of the candidate and their experience, and the idiosyncratic process that precedes the award of a PhD for those candidates who have 'done enough'. While the external examining system in existence for all UK higher education degrees extends to doctorates, the mediating role of the external in respect of a cohort examination that is intended to assure equivalent standards in taught degrees is prevented in the case of the PhD because each examination is unique. It is arguable that the range of attainment that exists above the pass/fail threshold is to be expected given the singularity of the PhD 'experience' and the different personal qualities candidates bring to their research. For some, however, this diversity of outcomes is undesirable. Variability is somewhat aggravated by the inconsistency in university regulations faced by examiners when deciding on their recommendations after the viva.

2.1 Diversity

By contrast with Lovitts' (2007) respondents, examiners in this study had as much to say about exceptional attainment as they did concerning lower achieving or borderline candidates, drawing comparisons that exemplified the diversity they perceived above the pass/fail threshold. A few also mentioned the 'middle of the road' candidate. My interview questions enquired about variability among candidates ('exceptional', 'good enough to pass' and 'borderline') but, unlike Lovitts' study, did not request categorisation of theses (dissertations) into broad categories (Lovitts used 'Outstanding', 'Very good', 'Acceptable' and 'Unacceptable'). Even so, respondents in my study acknowledged, both implicitly and explicitly, differences in attainment among candidates in their experience who had passed the examination, taking account of both thesis quality and viva performance. Comments ranged from straightforward statements, for example: '*Yeah, huge, huge variability in achievement.*' (STEM-E2); '*There is quite a bit of diversity... within the categories available...eg: exceptional/no corrections; minor corrections; major corrections/referral/fail*' (STEM-I1); and '*I'm constantly evaluating... whether it's excellent and just flying through and is intellectually of a very high level, conceptually but also in the execution and the presentation, or whether it's a good pass or a pass or not...*' (AHSS-I4), to more descriptive analysis, such as:

At the upper end, the work is extremely original and exciting; of global importance in the field. Very few students would achieve this. At the lower end, there are specialist studies that are using methods, analyses and interpretations already established in the field, to which a new study system is applied. (STEM-E1)

This observation differentiates the characteristics of both exceptional and unexceptional candidates and sets the context for the range of recommendations available.

Another external examiner made a comparison concerning a case I observed, using percentages of candidates in different bands:

The one that we examined, I would say was probably in the top third of PhDs I've examined, maybe between top 20% and top 30%, so not in the top 10% but just below that, but there is quite a bit of variability. (STEM-E3)

The top 10% of candidates was a measure also used by an examiner who cited the '*agenda-setting*' standard associated with the UK's Research Excellence Framework (REF) as a benchmark for exceptional candidates. Having suggested that only around 10% of the research papers published in a subject might meet the highest REF standard, they added: '*Some [candidates] will achieve this both by luck and by exceptional ability, but it's not an expectation of passing a PhD*' (EARTH-Ex1).

Viva observations and interview comments such as this enabled me to attempt to 'benchmark' the attainment of candidates in different fields. It appeared to me that two or three stood out as high achievers, while the rest more than met the pass/fail threshold, some at a higher level than others. The type and nature of corrections contributed to these comparisons, further evidence of which is in the case studies at Appendix 1.

Another examiner used only the lower end of the scale to describe candidate diversity:

There is variation in...quality...I've passed candidates I...consider marginal, usually because of a lack of breadth and/or depth of understanding in their subject. They may have done something quite narrow...which they can't necessarily put into a broader context. If they've done just enough work but aren't stunning you can pass them, e.g. [the thesis] is serviceable but not particularly interesting, or...they have found something interesting [but] they don't necessarily know why. This is different from those who haven't done enough work before submitting. (STEM-I1)

This response revealed the complexities of judging candidates whose work is unimpressive yet good enough to pass and the reasons why they might fall into this category. The examiner implicitly recognised that candidates who fell into any of these groups would be required to complete varying amounts of corrections in order for their thesis to become 'passable'.

Focusing on a wider range, another examiner discussed the *'broad divisions [between candidates] that you have to take into account when you first encounter the work'* (AHSS-E6), and suggested examiners would have an easier task if at the beginning it were possible to identify the candidate as exceptional *'because either the work or the methods or the argument are exceptional'*. However:

If the candidate is not exceptional [and] the work is...stodgy but...substantial...you pass them, because they've done the job, not because it's...shaking the discipline. (AHSS-E6).

The ability to 'shake the discipline' was a characteristic suggested by one or two other examiners to describe exceptional candidates whose research changed the way in which the discipline was regarded.

Some respondents based their comments on the pass/fail threshold, for example: *'I think [attainment levels] are quite broad. But...there's a huge amount of variation above [the] threshold'* (STEM-E4). This AHSS supervisor described the range in more detail:

There's...a baseline which everybody has to meet to get the [PhD]...Within that...there is a lot of diversity of extremely strong theses that...you could publish from, and ones that just about make it...because there's no grading criteria you've only got...a sense that there is differentiation between...different kinds of thesis. Sometimes, ones that just go above the line, it's not that the...work is not very good or that the person isn't very good, it's...that they've got to the point where they just want to get the thesis in and they've not gone for the most exceptional thesis on the planet...It's not like they're not capable of doing exceptional things, but...it's just that the [PhD] process exhausts people and they ...want to get it out of the way. (AHSS-S2)

In addition to recognising a wide range of attainment, similarly to STEM-E1, this supervisor was suggesting that doctoral work may be affected by the candidate's tenacity and staying power, potentially qualities that indicate 'doctorateness' (Denicolo and Park, 2010:2), or by their wish to move on to the next stage in their career. The possibility that some candidates whose work is considered borderline might in different circumstances be capable of producing an exceptional thesis is an interesting perspective relevant to other comments concerning the pressure on candidates to submit within a prescribed period during which they must also successfully complete structured training and develop a range of personal attributes, perhaps not allowing their full potential to develop.

An internal examiner in a STEM subject reiterated the broad range of achievement above the pass/fail border and contended that candidates would need to do extremely badly to fail at the final hurdle:

There obviously is a range of achievement...[it] is probably quite broad...The minimum required to...pass is much lower than a really good [candidate] will provide in terms of the amount of information and the amount of new stuff...At the lower end, the border between...a Masters...and a PhD thesis becomes quite thin and...close, while, at the top end, some students have done probably several PhDs' work by the time they've finished...The range is...broad [and] the threshold for failing is high...You have to do pretty badly...to fail...The default is not to fail somebody. (STEM-I3)

In alluding to the 'default' being not to fail the candidate and the high failure threshold, this examiner reflected a position that other respondents implicitly shared, even those who did not articulate it. In my observations, it appeared from the examiners' questions that they wished to test the candidate against standards in the discipline, for example, to assure themselves that the candidate understood the statistical methods used and that they were appropriate to the study. In most cases, examiners' feedback during vivas included improvements to the thesis, which could also be viewed as helping to maintain standards. None of the candidates whose viva I observed appeared to be at risk of failure, but a clear range of attainment was apparent and those 'at the lower end' received constructive advice to improve the final version of the thesis. Examiner STEM-I3 was one of two respondents to raise the related issue of using the standard of a masters thesis to benchmark lower attainment in the PhD, arguably a logical concept given that one of the alternative recommendations is to award a lower postgraduate degree. Yet the option available in many universities to award a lower degree to candidates who have not met the doctoral standard is rarely chosen.

Two AHSS internal examiners expressed concern about the amount of variation among (unobserved) candidates in their experience, who had passed the examination. One suggested that the 'unpredictability' (AHSS-I1) and lack of regulation of the process were factors leading to the diversity of candidate achievement above the pass/fail boundary. The other was troubled about the range of attainment in their subject, which they linked to the examination process and to the individuality of the thesis:

Yes...even within the passed vivas...there's quite a big range and that is a slight worry...more than a slight worry - that is a concern. But...that's the nature of the thesis, to a certain extent. (AHSS-I2)

An external examiner in the same subject, having been highly critical of the examination process, describing it as a 'crude system', elaborated:

It's higher stakes and it's very complicated because there's so many factors...[e.g.] a good student...who's been badly supervised and therefore turns in a not very good performance...When you talk about ability, it's a complex mix about opportunities and support and guidance. (AHSS-E1)

This respondent highlighted the capacity of other factors to increase the complexity of examiners' judgements. One or two others also mentioned inadequate supervision as a reason for some submitting an unsatisfactory thesis.

By contrast, two other AHSS examiners were much more relaxed about diversity. One internal examiner linked it to the wide range of students admitted to programmes: *'Yes, comparability and diversity...it's interesting because I do think that the PhD students we take are incredibly diverse'* (AHSS-I4). Another overtly suggested that diversity among successful candidates could have a positive impact on judgements, enabling examiners to compare one candidate with another to gain an overview of standards:

Diversity is as important as...comparability...you want to be sure that...they hit the right mark, but you don't want to be put off by the diversity, which is actually positive, by the ...differences between one candidate or another...in the practice PhD you can see why diversity is more of a problem than with a normal PhD [where] that diversity question doesn't arise really. (AHSS-E6)

While this examiner considered practice-based PhD outcomes particularly susceptible to variation compared with 'normal' PhDs they had examined, the data from this study suggest a complex picture that shows diversity exists across subjects and candidates.

The apparent acceptance, even expectation, by most examiners of the range of attainment of successful PhD candidates perhaps suggests that diverse outcomes are inevitable, especially given the impact of individual ability, background, circumstances and research context on each candidate's thesis and viva performance. The key point that most examiners and supervisors agree on is that in their subject they can identify a threshold, indicating that a candidate has 'done enough' to be awarded a PhD, once they have completed any corrections to their examiners' satisfaction. Examiner AHSS-E6 went as far as to suggest that diversity could be interpreted as a positive feature of PhD examination outcomes. While this was the only fully positive endorsement of diversity, the data suggest that only two respondents had any real concerns about different levels of attainment above the threshold and that some were prepared to assert a degree of comparability.

2.2 Comparability

The question of comparability of standards among UK universities has long been debated with no definitive outcome, in particular the question of whether a degree from University X is similar in standard to its equivalent from University Y. Such questions are sometimes met with answers that reflect the difficulties inherent in trying to compare universities with vastly different histories, missions and student profiles. Questions of comparability of standards extend to and are equally challenging for the PhD and its examination outcomes. While fewer respondents commented on comparability than diversity, three STEM examiners in particular made compelling arguments for why it was logical to expect *reasonable* comparability of outcomes in parallel with some diversity, even across disciplines, linking this position to the existence of generally high standards in their field and highlighting additional factors that can affect PhD outcomes.

Having acknowledged that their experience of PhD examining was 'limited' to a small number of universities, and that standards may differ, this examiner referred positively to a range of outcomes they had become aware of through colleagues' anecdotal references:

I imagine there's more of a range than I've [experienced], because...most of the people I've seen are really good...Maybe they're all good, I don't know, but I've never encountered a really bad [candidate], someone being entered into the process and not being able to defend themselves and their work, but I've heard about that from other people. (STEM-I2)

The implication here is that there is overall some comparability of attainment within and across this examiner's field, even though a few candidates fall below the required standard. The examiner also acknowledged that their experience was constrained by the relatively high attainment levels of the candidates they had examined. Another STEM respondent, with experience of examining across a range of scientific disciplines, agreed that standards were relatively high in their field, within a relatively narrow band of attainment and was even more confident about standards, expectations and comparability:

The standards are high overall...the minimum that a student needs to do to pass a PhD in the fields of biosciences and clinically-associated PhDs, is generally pretty high. (STEM-E5)

This examiner also provided a field-specific example of how more challenging research can affect thesis and examination outcomes, by describing how in their field (microbiology), some of the most challenging research with bacteria that is difficult to culture or manipulate genetically, might result in fewer data than might normally be expected, yet lead to a '*breakthrough that could open the field up*'. The response described challenges particular to

one field that may resonate in other disciplines and suggested that examiners adjust their expectations in cases where fewer than normal data generate significant results. The response suggested candidates should be supported in designing high risk and challenging studies if it might lead to ground-breaking results. More challenging research may take longer and the time periods specified for studentships are finite, as is clear from this extract from the Terms and Conditions of Research Council Training Grants: *'Doctoral studentships will range between three years and up to a maximum of four years full time support depending on the student, subject and Research Council expectations'* (UKRI, 2018). Another factor putting pressure on universities and candidates is that the criteria taken into account in REF judgements include candidate submission and completion rates.

A third STEM examiner made a convincing argument for comparability due to the process:

There is a level of comparability...because all...[candidates], regardless of discipline, will have spent three or four years of their time completely devoted to a research topic...if they've passed...it shows a level of commitment to that topic...if you have a PhD...you've shown a level of scholarship [in] that discipline...that is comparable across it...the nuances are differences in the...disciplines [which]...should be different...There is a level which...is probably generally met across the disciplines. (STEM-E2)

This respondent was supportive of the 'nuances' in different disciplines that are inevitable due to the academic and cultural contrasts among subjects and fields, yet they suggested a comparable standard as a result of similar experiences and training.

The following response from an experienced examiner summarises some of the key points concerning diversity and comparability, showing insight into the strengths and weaknesses of individuals and the range of attainment among those who pass:

There is a range of abilities. Some students find the viva enjoyable and afterwards there is little need for any...changes to the thesis. Even before they submit...these students are working at the right level...they will do well in the viva and tend not to have any corrections. You know this because their work is polished – it looks like the perfect PhD... They are good at arguing their position and have a good knowledge of the field...There are other students who are good at talking about their work but aren't at the highest standard and the thesis needs to be corrected; they are good in the viva at arguing their case but the written work is not as good as the best students. Then you have students whose written work isn't brilliant and their ability to discuss and argue their case is not as good as the others. They have a larger amount of corrections, e.g. they may be asked to re-write a chapter and may not have the ability to argue their case in the viva in the same way as other students. Their thesis is probably borderline but they will be able to defend it in some form. (AHSS-D1)

Comments from AHSS-D1 not only explore the diversity of candidate attainment, but emphasise the interdependency between the thesis and the viva. They suggest the best candidates first show their exceptionalism through the thesis and reinforce this judgement in the viva, that a second group present a less polished thesis but can make a strong argument in the viva, while for a third group, the thesis may be borderline and presage potentially significant corrections especially if a convincing case is not made in the viva.

STEM examiners appeared more confident of comparability and sanguine about variability among candidates who passed the examination. The majority appeared to accept a relatively broad range of attainment above the threshold standard within their field. One or two were confident about making cross-disciplinary comparisons. Unsurprisingly, it was rare to find respondents who had experience of examining PhD candidates across a range of subjects because of the specialist nature of academic practice. However, one examiner who had been involved in *'PhD examinations across the fields of microbiology, biochemistry and...chemistry'* was able to make the following comparison: *'I would say the expectations in those three [subjects], albeit in a fairly narrow set of fields, is pretty much the same'* (STEM-E5).

The data in this section show that the majority of respondents accepted the range of attainment above the pass/fail threshold in the PhD, that some expected it to be the norm, and that one or two even favoured such diversity. However, some would prefer a narrower range of attainment and a few were uncomfortable with the extent of diversity among those who passed the examination, leading to concern about standards. The introduction of grading in the PhD to enable some overt differentiation of attainment above the threshold might be considered a potential solution to the concerns regarding variability. Respondents were also asked to reflect on the range of specific examination outcomes available to examiners, using their own experiences.

3 Examination outcomes

While the different categories of corrections and variable time periods set by individual universities for the outcomes of doctoral examinations increase concerns over inconsistency and fairness, they nevertheless provide a flexible framework for examiners' judgements at the end of the viva. Respondents' experiences varied, especially regarding the proportions of candidates they had encountered who fell into different categories of initial outcome. At the lower end of attainment, some respondents provided an approximate percentage to represent the candidates who in their experience had failed, been awarded a lower degree, or been

asked to undertake major corrections, whereas others preferred to comment on the causes of failure or other less favourable recommendations emerging from some final examinations. Table 7.1 (p.190 and see Appendix 6, section 3) summarises the most common options open to examiners. While some outcomes are similar, the related time periods for each option are often different and may be interpreted slightly differently, depending on the regulations. In comparing universities' regulations for this study with respondents' experiences, it was clear that the category of minor corrections/amendments had potentially the widest range of interpretation and that examiners may have discretion to recommend this outcome in a variety of circumstances. On the other hand, some regulations are more specific, for example the University of the West of England (Appendix 6, section 3: K16.7.10), states that 'new work should not normally be required' if minor amendments are recommended. The following subsections explore the range of recommendations open to examiners and how they are used, beginning with minor corrections.

3.1 Minor corrections/amendments

Data suggest a wide range of outcomes exist within this, the most frequently used category of recommendation available to examiners. For example: '*Minor corrections can be anything between a few minor typos and re-writing some sections*' (STEM-I1). Candidates also highlighted the apparent flexibility attached to the recommendation for minor corrections suggested by this examiner.

The heterogeneity within the 'minor corrections' category, which several respondents suggested was the most frequent recommendation, is of particular interest. As suggested by STEM-I1, a recommendation of minor corrections or amendments ranges from almost insignificant changes to the re-writing of parts of the thesis, an observation borne out by the small number of examples in Appendix 6. Data suggest several factors contribute to this situation, for example, the tendency for examiners to support to successful completion candidates whose research they regard as having integrity and worth, even if the thesis is unimpressive. The extent of corrections recommended may also depend on the examiners' judgement of the importance of the contribution and the candidate's future plans, if they are aware of them. In this context, different views were identified. For example, one or two examiners suggested that if a candidate was not planning an academic career, they might recommend only amendments that would ensure the final thesis met an appropriate standard in the discipline, whereas if the candidate had already secured a post-doctoral role and might

be publishing, either from the thesis or separately, they were likely to recommend more extensive corrections. The following two STEM examiners, for example, were not in any way prepared to compromise the standard of the thesis, yet suggested they may adopt a differential approach to minor thesis corrections depending on the candidate's career path:

[Knowing the candidate's career plans] slightly modifies what you require...to progress, because if they're leaving academia, then is it really worthwhile pushing them into, say, improving their statistical approach? If they're not going to use that in the future, why are you wasting their time? (STEM-I2)

The second examiner agreed and summarised a differentiated approach:

If somebody has done a PhD and then...they're going out of [academia]...you can be a little bit more lenient because there's a pass level. If you look at the criteria of what actually constitutes a PhD...it has to be legible...it has to be a contribution, it has to be their work, and it's actually fairly limited...if you read the regulations. (STEM-E2)

Although against any formal framework or recognition of such differentiation, STEM-E2 compared candidates who were entering academic and non-academic careers to justify it:

It's [not] a bad thing to think...this person wants to be a high-class academic, and...if they're coming into that community, we have to make sure they are [that], versus this person is going off to do charity work [for example] and never pick up an academic journal again [even though] being Dr so-and-so might help.

They contended that this would affect their approach in the viva: *'It doesn't compromise the standard...you think they should meet, but it changes the rigour...and the examination of [the] content'*. They concluded: *'The standard is still equivalent, but what you do in that viva and how you approach it is a little bit different'* (STEM-E2). They emphasised that while permissible, this approach *'should never be a formal way or route to assessing'* (STEM-E2).

An AHSS examiner was equally committed to maintaining the standard but, in contrast with the two STEM examiners above, was sceptical of the desirability of asking any candidate with minor corrections aspiring to an academic career to undertake amendments which they perceived as wasting time on the thesis when it was publications that mattered:

You can ask the candidate to do quite a lot but I no longer think it's worth it at PhD level. However good the thesis...it's not going to be published in that form. You have to ask, does it meet the criteria? What would it take to make it presentable for archiving in a library or electronically? Any tidying-up can be done pre-publication. It's not good use of anyone's time to have a perfect thesis. (AHSS-E3).

One or two other respondents agreed with the essence of this examiner's argument, in that when a candidate was aiming for an academic career, it was more important to focus on publications arising from the thesis than on perfecting the thesis. In other cases, such as the

next example, where examiners' recommendations might be significant but not affect the findings, they might choose the 'minor' category, for example:

[One] example I remember is where a student had to complete a re-analysis but this was called minor corrections because the re-analysis did not change the findings – the result was consistent with the original conclusions. (STEM-E1)

Together with the examples in Table 7.1, respondents emphasised the inconsistencies inherent in the 'minor corrections' category. Most universities allow less time – one to three months – for the completion of minor corrections, and more, such as six to nine months, for major corrections, but there is no uniform approach. In some cases, examiners may use the relevant university's regulations to offer the candidate a different time period from that normally attached to minor or major corrections. For example, Appendix 6 shows that at Manchester University examiners who judge that '*more than four weeks of work*' is required to complete minor corrections can '*in exceptional circumstances*' recommend a period of 12 weeks instead of four for a candidate to complete these, '*based on the quantity of the work required and length of time [in which] it is feasible*'. If an extension to the period is not recommended by examiners '*In very exceptional circumstances, the candidate may apply to the appropriate School or Faculty...for permission to submit the corrected thesis later than the four...or 12 week deadline given*'. Similarly, examiners at Oxford University may recommend to the relevant divisional board an extension to the normal one month period for minor corrections. In such cases the board '*may, on good cause shown by the candidate, grant an extension of one further calendar month*'. The majority of universities make provision in their PhD assessment regulations for candidates to re-apply for extensions in mitigating circumstances. It is normal in some universities to specify that candidates may submit their revised thesis in advance of the deadline. Some candidates may find it realistic and advantageous to complete 'major' corrections in a shorter period, whereas others such as those already working or planning to take up a post-doctoral or other position, may require a longer period in which to complete 'minor' corrections, even if relatively unproblematic. The relative flexibility permitted to examiners and candidates as exemplified by regulations such as those at Appendix 6 is in some cases beneficial for candidates, even though it does not support calls for greater consistency.

The outcome for most candidates in this study was minor corrections. In a few cases it was clear that, although the amendments themselves were not negotiable, examiners were prepared to be flexible concerning time periods for completing corrections, depending on the candidate's circumstances, as long as university regulations permitted this.

3.2 Major corrections, referral and the award of a lower degree

Most examiners had little or no experience of candidates failing the final examination and not all respondents had experience of the award of a lower degree, yet most had experienced major corrections. Table 7.2 summarises the numerical responses provided by some, not all, respondents. A few preferred to give actual numbers, whereas others could not be so precise and therefore expressed proportions as percentages. Not all respondents chose to separate (a) the award of a lower degree from (b) recommendations of major corrections / referral, so numbers and percentages provided by those who gave combined responses are captured separately in (c). This variation in response choices is the reason for the mismatch between the total responding in the top row and the slightly larger totals if the column numbers are added up.

Examiners (Total: 7 AHSS, 7 STEM)		Supervisors (Total: 6 AHSS, 1 STEM)	
AHSS	STEM	AHSS	STEM
a) Had experienced award of a lower degree			
Less than 3%: $n = 1$	10%: $n = 1$	0: $n = 1$	
More than 10%: $n = 1$		Less than 3%: $n = 2$	
		More than 10%: $n = 1$	
b) Had experienced recommendation of major corrections / referral			
4-10%: $n = 1$	0: $n = 1$	10%: $n = 1$	
10%: $n = 1$	10-20%: $n = 2$	30%: $n = 1$	
	25%: $n = 1$	33%: $n = 1$	
c) Combined answers concerning award of lower degree and major corrections / referral			
Less than 3%: $n = 1$	0: $n = 2$	Less than 3%: $n = 1$	0: $n = 1$
4-10%: $n = 1$	4-10%: $n = 1$		
15-20%: $n = 2$			

Table 7.2: Examiners' and supervisors' experiences of candidates being awarded a lower degree or required to undertake major corrections (potentially including re-submission of the thesis)

While these numbers cannot be claimed to be representative, the variety of experiences is striking. It would be interesting to explore the frequency of different recommendations with larger numbers of examiners and supervisors to determine whether they hint at any differences among subjects or are affected by institutional regulations, for example, the range of options open to examiners that currently vary from one university to another, and the variety of related time periods accompanying within which the candidate must complete corrections.

One STEM examiner who had no experience of major corrections suggested they were rare because *'The definition of what the thesis should be is so broad that you can accommodate*

quite a bit through minor corrections'. The same examiner also suggested that, where examiners were aware the candidate may be intending to publish material from the thesis, they could *'get around'* recommending major corrections by making clear in their report that *'something wrong needs to be changed before publication'* (STEM-E1).

Highlighting another potential inconsistency, some understood that major corrections would involve a named examiner *'signing off'* the corrections, whereas others suggested that all instances of major corrections would involve the candidate resubmitting the thesis, for example: *'The distinction between minor and major corrections is that major corrections always involve a re-submission [of the thesis]'* (STEM-E1) and *'The majority of cases are going to be either minor corrections or...revise and resubmit but without a second viva'* (AHSS-E1). On the other hand there was agreement about the rarity of second vivas, mentioned in passing by one or two examiners. The question of differences between *'referral'* and *'major corrections'* was not an area of exploration for this study, but is of interest nevertheless.

As shown in Table 7.2, experience of recommending a lower postgraduate degree as a result of a PhD examination was limited, for example: *'an MPhil for a PhD is pretty rare'* (AHSS-E1). The recommendation of major corrections, on the other hand, was contended to be fairly frequent such as in this examiner's field involving experimental science:

Quite often, I'd say, there's major corrections...the question you have to ask is whether or not they need to do more experimentation and that's usually a big problem because they're no longer in the lab...where they did their thesis. (STEM-I3)

Having established that the practicalities of this situation mitigate against asking a candidate to undertake more experiments, even though they had occasionally helped someone to negotiate another year's lab experience, the examiner suggested a more frequent occurrence was asking a candidate to re-write their introduction because of a lack of general knowledge of the field and the need to present their data more effectively.

Responses concerning whether fewer or more candidates should fail the final examination or be awarded a lower degree were revealing. While most thought decisions they had been involved in had been correct, some added caveats based on anecdotal evidence that suggested occasionally passes are marginal, for example: *'[I'm aware of] some candidates...who have only just passed. But I haven't ever felt obliged to pass someone who shouldn't have done'* (STEM-I1). Another examiner from an observed case went further:

[If more should fail], it's only...marginal...It's not as if another 30% should fail or be awarded the lower degree. As you probably saw, I'm...on the tough side...there are people who are less tough, so I think that's what we're talking about...It's [not]...a quantum leap. It's probably...slightly on the too positive. (AHSS-E6)

Others suggested that, given progress hurdles in PhD programmes, no-one attempting the final examination should fail. For example, having confirmed they had heard of only one candidate in their field failing the examination and that had been *'almost 20 years ago'*, one examiner's view was that:

This may partly be because of the end of first year assessment that began to be introduced about then. At first, passing this was a formality but now it is a much more rigorous part of the process. Going through the upgrade can help the student realise that a PhD isn't for them. (STEM-E1)

In many universities progress hurdles occur not just at the end of the first year but more regularly throughout the PhD. Other examiners who supported the view that progress hurdles have an important role in preventing failure in the final examination included: *'No, I don't think [more should fail], because there's now very good procedures...that go through the thesis to rule out people...by the time you get to the end...it's reasonable that most people pass'* (STEM-I3); and *'There's more focus...you jump through the hoops at the right time and then you get to the end, and it's formulaic, and then you get your PhD'* (AHSS-I3). Others agreed that generally the correct decisions were made and that progress hurdles played an important part in 'weeding out' weaker students:

[Some] students...drop out...during the three years...because they are too weak and/or have failed the upgrade...If they meet the threshold...they're passing what we think they should...Students are being given the opportunity to defend and then correct their work. If they can't do this, they should fail. If something is wrong with the whole direction of their work, they should fail, but they should never have got to that stage. (AHSS-D1)

AHSS-D1 suggested that in their field, upgrades and other tests would mean only around eight or nine students out of an original cohort of 15 would progress to the final examination, citing the viva as a genuine potential point of failure, and emphasising the opportunity provided by the continuum of judgement. Others thought failure should not normally occur at the viva: *'Unless somebody's really tried very hard to not...do the bare minimum...most people wouldn't fail. I'd say it's probably about right.'* (AHSS-I3); and *'The student would normally be advised not to submit. By the time the student arrives at the final PhD examination, no-one should fail. If they do, it's a failure of the supervisor'* (STEM-E1). The role of the supervisor in getting the student to an appropriate standard is emphasised here and in several other contexts. A supervisor, who agreed that few students who progress as far as the final examination should

fail, used an example of referral to illustrate the support that would be provided to candidates in such circumstances:

Weak students, who perhaps should fail, are usually given a referral, and...such detailed feedback, provided...they follow that...they would have grounds for appeal if they were to then fail in the final examination...it's actually very difficult to fail. (AHSS-S1)

Another supervisor was unwavering in their confidence that standards were being upheld: *'In my experience, I don't think anybody has ever passed the final examination without me feeling they really deserved the PhD'*. Speaking from the perspective of both supervisor and examiner, they added: *'Internally or externally, I don't think we've ever compromised our standards because, if you make major corrections, you really make them re-write it...In one case, I think we re-viva-ed them as well'* (AHSS-S2).

A few were less confident about current standards, such as the next three respondents, the first suggesting, from *'hearsay'* (STEM-S1) that occasionally, candidates who had not reached a good enough standard were getting through. The second was unambiguous in stating:

'Probably more candidates should fail the final examination or be awarded a lower degree' (AHSS-S1). A STEM examiner agreed, situating their comments in an international context:

Maybe more should be awarded a lower degree. In terms of the standards of British PhDs compared to US or European ones, I think the [British] standard is slightly lower, and people get the degree and then often don't find appropriate work afterwards, so it may not have done them a favour, being awarded a PhD when a Masters might be more appropriate. (STEM-E3)

That UK PhD standards might be lower than those found internationally and that this might be reflected in subsequent employment opportunities, was an issue raised by this examiner only. It may relate to a particular field but would be a concern if a widespread view. An internal examiner, while not suggesting that more candidates should fail or that UK standards were lower than elsewhere, did believe that thesis quality might be affected as a result of *'the number of people being forced to submit before they're ready'* (AHSS-I3), a reference to the ever-increasing focus on timely submission.

An examiner in a practice-based STEM discipline emphasised their reluctance to fail candidates and referred to *'the in-betweens'*:

I am not a person who likes failing people...I don't think it's productive...On the other hand, it's unproductive to pass them if they're not worth it. That's why you have the in-betweens...if somebody gets major corrections, you give an opportunity to this person, and if these major corrections still require further major corrections, then I don't think you should pass them...but you always...in life, need to give a person a second chance. (STEM-I5)

One examiner focused on whether more candidates should be awarded a lower degree and suggested that *'if we re-organised things that enabled people to have that more as an option'* they would support more use of this recommendation:

Because...some people get trapped in the system...they want to get out...It's hard to get out because they've got to get on and...do their PhD...and if they leave, well then they've got nothing. It's tough...for some people it would be much better if, earlier on...they thought...`I'm going to take the quicker way out' – but...it's very rare that people do that. (STEM-I2).

This examiner's comments suggest that examiners are reluctant to use the option to award a lower degree: *'the whole system is designed to get people through and [other options] are not really enacted for students'*. Emphasising that examiners are reluctant to fail weaker candidates, they suggested that most want *'to get them up to standard'* and that the award of a lower degree was *'not implicit in the system'*. They concluded that one reason for the avoidance of failure and the award of an MPhil was seen as a failing by the university and a risk to funding streams: *'we receive money, from the university or from external people and they want to see this product...so...there's a huge pressure to get them through'* (STEM-I2), another reference to the pervasive emphasis on timely completion and the pressure on examiners, the supervisory team, the department and the university to push candidates who have made it through progression hurdles over the *'pass'* threshold in the final examination. Given that examiners and others are likely to take a view about the reputation and standing of a candidate's department, school and supervisors based on knowledge of graduate achievement, a record of failed candidates or output of low quality theses reflects on the training and support available in the department and on their admissions choices. This suggests it is in everyone's interests to maintain standards.

While the majority of respondents who commented on the categories of major corrections, referral and the award of a lower degree were confident of maintaining standards in their field, a few raised important considerations concerning marginal passes and the pressures on examiners to pass candidates. Some saw the award of a lower postgraduate degree as equivalent to failure, suggesting that STEM-I2's comments above concerning this option not being *'implicit in the system'* may resonate in other fields.

3.3 Failure

Data demonstrated that most examiners had rarely failed a candidate, i.e. only four out of the fourteen examiners in the left-hand columns of Table 7.3 had any experience of failing a candidate. AHSS supervisors who responded in respect of candidates they had encountered

when acting as supervisor or examiner were similarly unused to failure in the final examination. Given that other studies have concluded that examiners expect the thesis to pass and that failures at the point of final examination are few (Mullins and Kiley, 2002; Pearce, 2005; Golding, Sharmini and Lazarovitch, 2014), these responses were unsurprising. Some were able to estimate the numbers in percentage terms, based on their own experiences, others gave a number rather than a percentage.

Examiners (Total: 7 AHSS, 7 STEM)		Supervisors (Total: 4 AHSS, 0 STEM)	
AHSS	STEM	AHSS	STEM
0 failures: $n = 4$	0 failures: $n = 6$	0 failures: $n = 2$	
2 failures in 20 years: $n = 1$	1 failure in 18 exams: $n = 1$		
Less than 3% failure: $n = 2$		Less than 3% failure: $n = 2$	

Table 7.3: Examiners' and supervisors' experiences of candidates failing the examination

A few respondents qualified their numerical estimate with a rationale for the low incidence of failure, or the reasons for submission of an inadequate thesis. Some re-emphasised the numerical response, such as: *'Real fails are very, very rare'* (AHSS-E1), and *'It's more likely that somebody is going to drop out than fail'* (AHSS-I4). Others took the opportunity to explore some of the reasons for the low failure rate (including progress assessments) and how candidates can be supported to raise their thesis to pass level, for example:

Examiners are very...unwilling to fail a candidate. Firstly, examiners believe that failure should have been sorted out by the university earlier on...so they feel that...if there's a shoddy product...[it]...is, at least in part, the university's responsibility...In extreme cases...examiners will complain to the university and say 'I don't think this person was supervised adequately'...Secondly, examiners often feel somehow...sympathy for a student, that they've spent four or five years of their life on this, and how could they fail at the last hurdle. (AHSS-S1)

The implicit reference to progress hurdles by AHSS-S1 supported comments by STEM-E1, AHSS-I3 and AHSS-D1 above, some of whom considered that candidates should not progress to the final examination if their work was not of a sufficiently high standard. They accentuate claims made by other respondents that examiners are extremely reluctant to fail PhD candidates for a number of reasons, not least because of the time and effort they have expended during their registration. One examiner, for instance, chose to relate this point to the financial commitment of international students: *'No-one wants to tell someone they've failed after four years' work and payment of £50k-£60k in fees. Overseas students are paying a vast amount of money'* (AHSS-E3). This response again emphasises the pressure on examiners and universities not to fail candidates once they reach the final examination.

This STEM examiner, on the other hand, focused entirely on the candidate's ability to obtain convincing research results:

The most common reason for failing is not...around originality but about the quality of results...if you don't believe the experiments, if they're designed such that they're not sound, or, even worse, if...they've been somehow faked...that's...definitely sufficient to fail...for the majority...Failing on the originality criterion, although...very important, is, in my experience, pretty rare...It would have to be pretty unoriginal...unless you copy...in biology, it's very unlikely that you will have done exactly the same...as somebody else. (STEM-I3)

It is notable that this examiner chose to contextualise instances of failure by mentioning the possibility of failure due to lack of original research, suggesting that for some the requirement for originality remains central to PhD assessment judgements.

The low incidence of failure indicated by respondents in this study supports the findings of other researchers such as Mullins and Kiley (2002) and Johnston (1997), whose results showed reluctance by Australian examiners to fail students whose thesis was poor. Even in the 1960s in the United States, it was rare to experience a fail at the viva stage:

At that point, at most...universities, it is extremely rare for a candidate to fail – not, as academic agreement has it, that many may not deserve to fail, but (1) because it is then too late for a faculty to assert itself "in all fairness to the candidate" and (2) because even though it may be feasible to fail the candidate, it is difficult or highly embarrassing at that point for a department to fail his sponsor, his committee, or even itself in the process.' (Berelson, 1960:200)

This quotation is from a critique of the thesis defence based on the views of a group of graduate deans. It underlines the importance of early intervention by faculty in the case of weak candidates, which in the UK is now largely achieved through regular progress assessments. The quotation also illustrates the reasons why, if such candidates have erroneously been allowed to progress, there is moral pressure on examiners not to fail them outright, especially given the personal and financial investment in the degree that extends beyond the candidate to supervisors and sponsors. The current study highlights the role of the viva in providing an opportunity for examiners to explore with the candidate the possible reasons for an unsatisfactory thesis, enabling them either to satisfy themselves that they should award the PhD, or to recommend major corrections or referral to address weaknesses. This reduces the pressure to pass underperforming candidates while providing an opportunity to maintain standards.

Given the range of achievement among successful candidates, I asked respondents to consider the benefits and disadvantages of introducing grading or a distinction to acknowledge exceptional attainment.

4 Grading and distinctions

There is no evidence for why, 30 years ago, UK universities disregarded an opportunity to adopt grading for PhDs following strong encouragement from a higher education sector-wide body (CVCP, 1988). It is in some senses surprising that the final assessment outcome for the PhD remains either pass, award of a lower degree or fail, especially given the acknowledged wide range of attainment among candidates and the graded approach for other UK higher education qualifications. Lower postgraduate awards such as Masters degrees, even if not graded, normally offer an option of 'distinction' or 'commendation', whereas the PhD does not.

This section addresses the question of whether the UK PhD should be graded rather than examined as a pass/fail degree, and if graded outcomes would be beneficial or of no advantage. Sections 4.1 and 4.2 explore the evidence against and in favour of grading and distinctions. Data suggest considerable agreement around the question of grading. While some respondents approved of it, even those with experience of examining in other countries mainly favoured the UK system. Similarly there was limited enthusiasm for introducing awards such as distinction for exceptional performance in the PhD examination, although a few were in favour of some kind of differentiation.

4.1 Arguments against grades and distinctions

The majority of examiners, supervisors and candidates rejected the idea of introducing any form of grading system for the PhD and made several different arguments, mainly based on their own experience. Arguments against moving to a graded judgement include: that it would increase subjectivity and emphasise individual examiners' preferences; that achieving a PhD is about going through a process and acquiring a set of skills which are best assessed by a threshold judgement; that grading could potentially lead to grade inflation; and that PhD examination outcomes do not lend themselves to graded judgements because each is based on a unique set of underlying factors and experiences. For example, having stated that grading would be unhelpful, one examiner suggested: *'Every PhD is so different. You would be comparing apples and oranges a lot of the time'* and added *'There is always subjectivity in the assessment but breaking down the judgement into fine grades would be difficult'* (STEM-I1).

Another examiner asserted *'I don't think we should grade PhDs'*, the rationale being *'There are so many factors which vary from project to project...candidate to candidate...supervisor to supervisor and [institution to institution] and: 'Every PhD needs to be treated on its own merit as a piece of work, with a unique candidate'*. This examiner suggested it was preferable not to grade the outcome but necessary to ensure *'[the thesis] reaches the minimum threshold for award of the degree'* (STEM-E5). Other robust views against grading included: *'No, I don't...see the need for it...it's not necessary. I'm not aware there is a demand out there for the grading of PhDs'* (AHSS-E3). *'I would be very sceptical about trying to grade within the pass...[the PhD is] not amenable to that sort of grading in any reliable sense'* (AHSS-I1); and

My gut answer would be to say no because [PhDs] are all so different. It's not a standardised test...it's comparing chalk and cheese in some cases so I can't see how you could grade it and maintain standards. (ENG-Ex1)

The same examiner also referred to grade boundaries that would be likely to differ between institutions, raising questions about comparable standards across universities and academics (Lovitts, 2007).

An AHSS convenor was unequivocal about the clarity of current procedures, arguing:

There is a system and the candidate can get an MPhil instead of a PhD. Grading would be quite tricky. It's bad enough having to pass or fail people on the basis of "Have they made a contribution to knowledge?" and "Could this be published?" These criteria are quite clear. (AHSS-IC1)

One examiner emphasised the value of the PhD in the wider contexts: *'A score for the PhD would be limiting – you're not only looking for academic excellence'*, relating this to the additional training some candidates embark on in preparation for their future careers. They suggested that would make it difficult to apply a fair grading system because of the different experiences of individual candidates. They also mentioned the indeterminate circumstances of PhD training: *'All kinds of criteria feed into how good a PhD might have been – a lot of luck is involved and it also depends on resources and getting the right project'* (BIO-Ex2).

Several examiners likened the pass/fail judgement to other degrees that confer a *'licence to practice'*. Two compared it to qualifying as a medical doctor, the first stating: *'You can't award degrees of PhDs. You can't be a "doctor but I'm a very good one" or a "doctor and I'm an OK one", or a "doctor and you're an exceptional one". That comes afterwards'* (STEM-E2). The second examiner in an AHSS field, focused on the PhD as a *'qualifying'* award and suggested:

Whatever they do later on, in terms of publication, is another story altogether...But if you look at [it as]...a kind of qualifying stage...that's an argument against grading...a PhD is a PhD...a medical doctor is not graded...you're now a researcher and you're allowed to practise as an academic, rather than it's an A, B or C. (AHSS-E6)

A supervisor with knowledge of the performing arts compared being awarded a PhD to winning a prestigious piano competition:

It's a bit similar to...piano competitions, opera competitions...wonderfully exciting things, but then it does happen that Van Cliburn wins the Tchaikovsky in Moscow, is celebrated around the world for five years and then disappears for the rest of his life...It doesn't mean anything to have won the competition in the long run. (AHSS-D2)

The arguments for retaining a pass/fail outcome support the idea that conferment of the PhD is a 'licence to practise' and that graduates prove themselves by what they achieve subsequently (and see Mullins and Kiley, 2002; Lovitts, 2006). Examiners in different fields agreed that the proof of the individual's ability came later, '*out in the big wide world*' (STEM-E2) when applying their knowledge and suggested there were arguments for and against grading the PhD, the thesis being '*not the first great work of a future academic but more like a driving licence*' (AHSS-E6).

Another examiner suggested grading '*could be open to abuse because...the important thing is whether the candidate's got the PhD or not*' (STEM-E5), adding that aiming for a high grade could distract able candidates aiming for academic careers from focusing on career-enhancing publications, a view shared by others:

The most important thing for a student in our field is to get papers out as soon as possible...a grading system could be a hindrance...The student might concentrate on polishing the thesis to get a higher grade rather than getting out some good papers, which are judgement by the scientific community at large, rather than two examiners. (STEM-E1)

The emphasis here on should peer review by a wider audience highlights the relevance of publications to candidates in disciplines where it is expected or permitted that prior publications will be included in the thesis. Accentuating the role of publications in scientific disciplines, one STEM examiner asserted: '*In science, you're judged far more on your publications which...are easy to quantify. So to some extent...the PhD is a process and a set of skills you've acquired and the yes/no is fine*' (STEM-E3). This view of the PhD as a process was shared by candidates and one or two other respondents. Another STEM examiner concurred with the prioritisation of publications over the outcome of the PhD: '*Coming back to the question of published work, if you've published three or four papers then probably, irrespective*

of what it says on your PhD...it wouldn't make much difference' (STEM-I3). The view that candidates are able to make their mark through publications was shared by AHSS respondents one of whom focused on the need to submit the thesis *'then make the exceptional contribution through the publications you produce'* (AHSS-S2).

Another examiner opposed to grading referenced the increasing inflation of A level grades to make the point that a similar situation could occur in the PhD, suggesting that grading could lead to too many candidates achieving a top grade *'and it then becomes...part of a qualification process, rather than indicating that here is a person who [represents] the highest level of autonomous work...in their own discipline'* (ECON-Ex1). The other non-case related examiner in economics made a similar argument.

The focus on publications, preference for a pass/fail judgement, and references to a licence to practice are all rooted in academic contexts and suggest affinity with the 'apprentice' model in which the PhD originated. Few respondents overtly acknowledged other careers, a somewhat surprising finding given that careers outwith academia are chosen by the majority of PhD graduates who, even if they aspire to academic jobs, face increasing competition even for post-doctoral positions. In this context, examiners suggested that academics rely on references or word of mouth, for example: *'the best students...stay on to do postdocs'* and therefore grading the PhD *'I'm not sure...would make a huge deal of difference'* (STEM-I3); and *'You can see who the supervisors were, the research group that they've come from'* (STEM-E2). These responses suggested that grading is unnecessary in the context of post-doctoral research because account can be taken of other factors that reveal the likely quality of the candidate:

If they're going to go into academia...the strength of their thesis gets reflected in the reference...I would write...'exceptional thesis' or 'exceptional candidate'. So, the fact that they've got a PhD and there's no grade to it doesn't...matter in the next step of their career trajectory...it's the reference they get from their supervisor that makes the difference. (AHSS-S2)

Several respondents concurred with AHSS-S2 when describing exceptional attainment, including this examiner: *'They will get the recognition if their work is so distinguished that it results in academic outputs arising from the thesis...peer analysis comes afterwards and...peer recognition'* (AHSS-I1). Another examiner took a direct approach to the PhD as preparation for an academic career to argue against grading:

Now, the PhD is only really a pre-requisite for an academic job. You can sort out the quality of candidates for academic jobs without their PhD being graded. Who needs this? The market will sort it out. (AHSS-E3)

Some respondents used the example of proving excellence through publications to refute the need for grading the PhD. In STEM, exceptionality would be apparent at the time of the examination because of the expectation in many science subjects of publication during training, whereas in AHSS, publications normally occur post-examination.

In addition to the medical doctor analogies, several examiners including AHSS-E6 referenced different assessment systems when arguing against grading the PhD. Two examiners chose GCSE and A level education for comparison and although neither was in favour of grading, they had different opinions about the systems in secondary education. The first, having confirmed they were not in favour of grading despite a great deal of variability: *'You pass or you don't. But there's huge variation in that top end'*, added: *'Almost the difference between a C at GCSE and an A*; and a C at A Level and a top A... They're all passes, but they're [at] various levels...and we only have pass or not'* (STEM-E2).

The second examiner was also reluctant to introduce grading: *'I haven't got a lot of patience for adding A*s and A**s at GCSE and A level - I'm quite happy with it being pass/fail'* (AHSS-E1). This examiner put forward an interesting argument against grading the PhD, suggesting that even if a 20-point scale were introduced, *'different examiners would use the scale wildly differently'* and there would remain no way of *'standardising'* judgements. They were of the view that the PhD *'is the...pinnacle of academic achievement'* and therefore to introduce grading would devalue the achievement of some graduates: *'Part of the culture should be to make sure that someone feels really pleased with themselves when they get their PhD'*. They suggested it was important to ensure that students who are required to undertake minor or major corrections *'don't go away with any sense of failure'* (AHSS-E1).

One examiner suggested that, in an international market, grading might be misunderstood: *'Many PhD [graduates]...go to America for their postdoc, or to Europe and I'm not sure people would understand the gradation anyway...I'm not sure how helpful that would be'* (STEM-I3). Another, having strongly rejected the idea of grading, also pointed out that *'It would put one more burden on examiners'* (AHSS-E3), an argument made by others.

The focus group, in particular FG-F1, implied that the categories of recommendation open to examiners at the end of the viva were in effect a form of grading, an approach supported by others, for example: *'The categories...either no corrections, minor corrections, major corrections, resubmission...there's kind of flexibility built into those criteria'* (AHSS-I4); and

'They're kind of graded...we have to grade them anyway in terms of whether it's a pass with minor [or major corrections]' (AHSS-I3), even though recommendations are not visible once the candidate has received their PhD. Another examiner agreed *'It works quite well as a threshold judgement...the various...options open to examiners are good'* (AHSS-I4).

Several respondents commented on the diversity inherent in academic disciplines as a challenge to any kind of grading:

If you've got a threshold which is essentially the PhD, with one or two fall-back thresholds, which the candidate must meet...that's a struggle in itself, [but] to find comparability amongst diverse disciplines, huge variety of methodologies, many different examiners...I don't think [the PhD] is amenable to more tightly framed mark schemes because each piece of work is [so]...individual...I can't see how you could be any more specific about the things you're looking for. (AHSS-I1)

Another AHSS examiner also focused on thesis diversity and potential subjectivity in arguing against grading:

I don't think [you can] grade it because the PhD is such a range of different types of document...the examining process is somewhat subjective and individualistic...ready or not is hard enough really. (AHSS-I2)

A supervisor in the same case agreed that differences among theses would make comparison difficult, yet suggested that as an examiner *'you must have a set of criteria in your head'* (AHSS-S2) to identify exceptional work.

A STEM examiner suggested that any kind of grading system for the PhD would need to evaluate supervisor input as well as the quality of the candidate: *'The supervisor would have to be more closely monitored...and...would have to give a declaration of...what help they'd given'*, concluding that *'every supervisor is different on what is the optimum'* (STEM-E4).

Others considered the outcome of a pass with no corrections an implicit grading or distinction, for example: *'In a sense there is...an implicit rating...If a student passes with no corrections, that is something they can and often do put on their CV. "Passed with no corrections" – ...it's pretty impressive'* (AHSS-I4). Two STEM examiners who were both against any kind of grading concurred with this view. The first suggested that including in a CV the statement *'Passed with no corrections'* (STEM-E2) would indicate exceptional achievement to any employer, while the other proposed: *'You can still say, "Oh, I got minor corrections" [or] "I passed with no corrections"...if you want to'* (STEM-E4).

While the opposition to introducing grades was fairly consistent, the approach to distinctions or prizes was more eclectic. Nevertheless, the majority continued to resist any form of differentiation, for example: *'I'm not in favour of a starred PhD. I'd prefer to stick to a threshold judgement'* (AHSS-E3). Another AHSS examiner regarded some kind of distinction for exceptional candidates, if embedded in the examination, as undesirable: *'If a university wants to run its own special prize for people who are judged to have submitted the best PhD of the year...that's up to [them], but as an institutionalised part of the examination process, no, why?'* (AHSS-I4). The option of institutional prizes was raised by two other examiners, one of whom referred to an annual *'Faculty award'* (AHSS-I3), but did not evaluate the usefulness of such prizes. The other, having confessed *'I'm not a massive believer in prizes'* (STEM-I3), then suggested *'One possibility would be [for] each university [to have] some sort of prize for the best thesis or something like that'*, adding that *'it's just another criterion for distinguishing people, [a] sort of box...but it's a whole lot more work'* (STEM-I3). An AHSS examiner remembered a *'best thesis of the year'* award sponsored by a disciplinary association, as *'one way of recognising distinction'*, yet did not ultimately support the idea: *'Take Einstein... everybody would have known Einstein was pretty exceptional and does he need a prize...no... not really, no'* (AHSS-S2). Three non-case related examiners commented on prizes. ENG-Ex1 recalled their recent experience as an external examiner, where they had been asked if they wished to recommend the candidate for a university award or prize. They did not object to this as an option but did not believe in a distinction noted on the degree certificate. The second suggested *'You should not get prizes for the thesis'* (BIO-Ex2), reasoning that successful candidates would have other opportunities for distinction. The third agreed: *'I would be against merit awards. I don't think it's necessary'* (ECON-Ex1). On the topic of distinctions, a STEM examiner returned to the question of differences between subjects and institutions, contending that *'it would be very hard to come up with a set of criteria that applied equally across disciplines'*, questioning whether *'a distinction from [university a] is the same as a distinction from [university b]'* and suggesting that *'We totally fail to do that at undergraduate level'*. In conclusion they summarised the practical challenges of introducing some kind of merit or distinction:

It's not just...did someone get a first, it's a first from where?...One thing I like about...[the] PhD is that the institution you've done it at...becomes less relevant. It's who you've worked with that's...important...You won't get uniformity so...in a way, it just becomes meaningless and...it gives another thing for the examiners to...quibble about...In terms of corrections, what do you do? Do you get them to correct it to...the merit level or just the pass level, or do you then have to re-grade it after corrections...What if the candidate really wants a distinction so they ask you to make corrections to a distinction level? The whole thing...would be immensely complicated. (STEM-E4)

These comments not only accentuate the practical difficulties of implementing any kind of merit or distinction (STEM-E4), they also draw attention to academic diversity (AHSS-I1) and the problem of subjectivity (AHSS-I2), a long-standing criticism of PhD assessment. Examiner AHSS-I1 made some positive observations about candidate attainment in concluding the discussion concerning grading: *'Credit should be given in all the usual ways...my...feeling is anybody who has achieved a doctorate has achieved hugely anyway and that they deserve congratulations for [it]'* (AHSS-I1). This examiner described the *'huge academic challenge which people undertake'* and, similarly to AHSS-E1, suggested *'they deserve the biggest possible congratulations for just passing'* (AHSS-I1).

Not all respondents were against distinguishing exceptional candidates from the majority who pass. Some made strong arguments for differentiation, as demonstrated in the next section.

4.2 Arguments in support of grading and distinctions

Although most respondents in this study eschewed any grading for the PhD other than the current range of post-examination recommendations a few, especially those with experience of doctoral examining in other countries, were prepared to support some kind of grade or distinction for exceptional candidates. Some, like this STEM examiner, for example, were cautiously in favour of distinctions: *'At the top end, I wouldn't be against some kind of distinction'* (STEM-E1) but qualified this by adding *'A student like that will be getting good papers out anyway. I'd prefer a continental system where the thesis is a set of bound papers plus an explanatory commentary'* (STEM-E1), referring back to the notion of peer reviewed publications being effective indicators of quality and reflecting subject conventions for thesis models.

An AHSS examiner was more positive about recognizing high achievement: *'The fact that someone has done first class work is something to be noted...we want this'* (AHSS-E6). They observed the importance of the distinction grade in the MA *'because that's a way of judging who is going to be succeeding at PhD level'* (AHSS-E6) an interesting point to make, given that the PhD's eminence as the highest academic doctoral qualification means it is not used to determine the candidate's suitability for another, higher degree. This examiner concluded:

You come to [the] PhD and everyone is the same. There's a...greyness of scale which is odd because...people are not the same...but you are not able to grade them...short of an A, B, C [which is] not very useful at that level because...all the candidates are above a certain level that you take for granted...just some method of separating the very exceptional would be...very positive. (AHSS-E6)

Similarly to other respondents, examiner AHSS-E6 did not favour a simple grading system, but was enthusiastic about some tangible recognition of exceptionalism in the PhD. An AHSS supervisor also supported some form of differentiation:

I...would quite like, just a star, a sort of asterisk for excellence, or exceptional, because...there is a huge difference. I know it's a threshold, pass/fail but...there are problems at the borderline. That's the most difficult...judgement...Exceptional performance in the final assessment...ideally for me would be rewarded simply with an asterisk...if it were recorded on the degree certificate. (AHSS-S1)

This supervisor added that it might be beneficial to candidates to be able to indicate on their CV that they had a starred PhD, and that they made sure any reference reflected exceptional ability. They cited examples of graduates who named their examiners when giving their thesis title in applications for academic jobs. A convenor suggested *'It might be possible to have a distinction for an outstanding piece of work'* adding that a recommendation to *'a higher body'* would *'take the decision out of the immediate examination arena'* (AHSS-IC1).

A STEM supervisor, who had *'been viva-ed in other countries'* and had experience of being awarded marks at the end of the examination, whose name suggested a non-UK nationality (but who had not examined internationally) stated *'Personally, yes, I like grades'*. Making an argument linked with the variability of candidate attainment they continued:

I...like to put...a measure on things and I have read PhDs that are very, very good and some that are less good and could have been better...they are all passes but it would be nice to reward or...to point out the differences. (STEM-S1)

This supervisor confirmed they would welcome a form of merit such as the *'cum laude'*, or *'with distinction'* system operated in some other countries, that would appear on the graduate's degree certificate. This option was supported by a STEM internal examiner: *'It's a mistake in the UK system that they don't have some kind of cum laude prize at the end...that would...be a good thing, to acknowledge...an exceptional piece of work'*. However, the same examiner qualified this view by suggesting that *'On the other hand, getting too much into fine grades would be idiotic as well. I think some kind of acknowledgement where things are exceptional would be sensible'* (STEM-I2). Another STEM examiner, also experienced in other systems, favoured grading, especially to reflect the contribution made by candidates who unequivocally led the research:

PhD students should be graded for the contribution and the effort they did to the study...where they actually own the PhD and develop [it]...That doesn't come through [in the present pass/fail judgement]. It [shows] in the future career because usually these are the people who are successful, but...it should have a grading element as well. (STEM-I5)

Others were brief in their support for some form of differentiation. For example, having clearly stated *'I think there should be a pass and there should be a distinction'* (STEM-I5), this examiner, similarly to some of those quoted above, declined to support a simple grading system: *'You cannot have A, B, C, D'*. Instead they argued for a distinction by describing a candidate who by adapting a particular technique created a benchmark for others in the field, who they suggested should be valued more highly than candidates who used the technique subsequently.

Candidates in this study were also asked for their views about the pass/fail judgement in the PhD and the possibilities of grading the degree.

4.3 Candidates' views about grading and distinctions

The range of views expressed by candidates about grading and distinctions showed striking similarity to those of other respondents and also fell into two main categories, against and for grading, with several open to the possibility of some kind of merit. Candidates also presented one or two additional arguments on both sides.

The majority did not believe that introducing a grading system would add anything to their attainment, such as this STEM candidate: *'I think that a pass or fail system is quite straightforward, so it's OK'* (STEM-C2). Another, part-time STEM candidate in a practice-based discipline agreed: *'I don't think it should be graded...if you have the quality to finish a PhD...they are more or less the same'* (STEM-C5), suggesting they perceived relative equivalence among graduates. Similarly to some others, STEM-C5 emphasized the difference between the PhD and other higher education awards, drawing attention to the experience and the process: *'a lot of things...have happened to you during three or four years and everybody has very different scenarios...the ease of getting...data and results'* (STEM-C5). This candidate could not see how any grading system could be fair and also disagreed with any *'congratulatory letter or prize'* (STEM-C5) being linked to the final assessment, although suggested that *'If [any] exceptional performance certificate or award [were] given'* it should be *'on the quality of the [candidate's] whole work and publications...and the difference in contribution to science'* (STEM-C5), extending the idea of recognition of excellence to a larger body of work than the PhD.

Another STEM candidate had a similar perspective on grading to STEM-C5. Having made it clear they thought grading would be unhelpful: *'I don't really know how you would do that because...if you grade it...you need to have some standards against which you can judge it, and I think that's very hard'*, this candidate explored the diversity of theses and research approaches: *'The theses that I...read [were] so...different...in length and number of chapters...I did fieldwork for a year which is completely different [from] if you were doing a library based one'* (STEM-C1).

Linking the question of grading to the 'rite of passage' argument, the same candidate who was on the point of taking up a post-doctoral position, observed that the pass/fail system was eminently suitable for the transition point between being a student and a qualified researcher: *'I...see this viva...like a rite of passage...you're finally an academic...once you pass you become one of them, whereas if they give you a mark it's more like a student relationship still'* (STEM-C1). The candidate supported the practice of a learned society connected with their subject that awards a medal biennially for the best PhD thesis: *'Within the field, you have these...awards, which I think...is very nice'*. They also suggested that the option of 'no corrections' to the thesis was similar on a CV to a kind of distinction and that *'If you have a PhD by publication...you are very likely not to get any corrections, whereas...if you do a normal PhD there will always be a couple of spelling mistakes'*. The candidate agreed with others that once the PhD had been awarded it was impossible to know the level of corrections: *'you always get three months for [minor] corrections, so it doesn't really show you if you needed to do a new analysis or if it was just changing a few spelling mistakes'* (STEM-C1). This candidate's views were shared by other respondents and introduce another dimension concerning the potential for different thesis models to affect a candidate's level of corrections.

Similarly to other respondents, an AHSS candidate saw the pass/fail outcome as appropriate for the PhD:

Instinctively I think 'No' and...that's probably because I see it as a hurdle...a single bar...either you've got it or you haven't...you either get it or you don't. There's no such thing as a better PhD than anyone else's, or if there is there shouldn't be. It's a...threshold...and anything over that is irrelevant. (AHSS-C1)

This candidate was aware of the existence of thesis prizes but contended: *'The congratulatory letter or whatever...wouldn't work'* and *'because all the examiners are different'* they would have no idea of whether a thesis *'was any better than others in the department unless they sat down and read them all, which is how...other prizes...work'* (AHSS-C1).

Another AHSS candidate was also against grading: *'I don't think it should be graded... academia is competitive enough already. It's individualistic, it's competitive'* (AHSS-C4). This candidate suggested that those aspiring to academic careers would potentially be at a disadvantage if some kind of grading system were introduced: *'The idea that you might go into academia with...let's say a B or a C...and if you've got a C you're not going to get a sniff at the door anyway'*. They acknowledged the current system did not provide any form of differentiation: *'You've got a PhD, but was it a good PhD or a bad PhD? Well, nobody will know, will they?'* but suggested other indicators were sufficient: *'Somebody who's done a reasonably good PhD is going to have the basis...a springboard to develop their work if they want to go into academia...whereas with a mediocre PhD you don't have as much of a springboard, so I guess what's in the PhD kind of talks for itself...I'm glad...it isn't graded'* (AHSS-C4)

Another AHSS candidate had a similar approach to STEM-C5, in that they could not envisage how differentiation would be achieved: *'How would you have a PhD with honours, or PhD starred, PhD plus? I don't think so, no...[it] should be just the plain PhD...either you've got it or not'* (AHSS-C3). This candidate, however, was not opposed to *'a separate distinction'*.

Only one candidate was prepared to consider the benefits of grading, but was unsure about how this might work in practice. Having acknowledged that *'there's a big difference between the two top bands that...exist now, which is "no corrections" and "minor corrections"'* and that *'There's a huge gap because there [are] always going to be corrections of some sort in a...body of work like [the PhD]'*, was not averse to some kind of grading:

If it was broken down a bit more or graded you could still have minor corrections and...you could have then a perfect [thesis], completely no corrections required...then...have something in between those two categories. (STEM-C4).

They argued that even minor imperfections such as: *'typos'* or graphs that *'don't make enough sense'* (STEM-C4) were still corrections. They supported recording *'no corrections'* in some way on the degree certificate, denoting a rare achievement.

The final STEM candidate, who had had experience of higher education elsewhere in Europe, was ambivalent about grading, but suggested *'it might be good to have differentiation between certain types of PhDs'* (STEM-C3). They explained that they had shared *'seminars and upgrade talks'* with PhD candidates in cognate but different subjects and had found some of them *'quite limited in their critical thinking, the design of their studies and how to take forward*

the evolving PhD studies from after the initial study'. This candidate was therefore in favour of differentiation:

It might be quite nice if there could be a distinction but I don't think it should be graded, as in having lots of different grade points because after all it's work of three or four years ...I don't necessarily think you could put a grade on it but maybe something like a *cum laude* or distinction' (STEM-C3)

This candidate acknowledged that the PhD was about the process of doing research and that one of the arguments against grading was that it was not possible to know to what extent the candidate was *'the problem-solver'*, or had just followed their supervisor's advice, suggesting that:

The student might not have had that much critical thinking or involvement in the study design or the process...in a PhD in science at least, when you're doing research it's more about the process than the outcome. (STEM-C3)

In concluding, this candidate was in favour of a distinction, but on the basis of the thesis only, for *'exceptional performance'*. STEM-C3 gave an unusual and very interesting perspective on the question of grading and distinctions, rooted in their subject. Their view about receiving a distinction for the research undertaken mirrors the comment of candidate STEM-C5 above in respect of the scientific contribution and supports the idea of a prize for the written work only.

As in other parts of this study, candidate responses show insight into some of the arguments integral to the questions concerning grading and distinctions and are generally in alignment with those of examiners and supervisors, especially in their opposition to the introduction of any simple grading system for the PhD examination.

5 Conclusion

The three distinct but related conceptual areas explored in this chapter present ongoing challenges for the PhD and its examiners. The first shows the tensions inherent between the diversity and comparability of PhD candidates, which nevertheless do not necessarily constitute, at least for many, a threat to standards. The majority, even those who did not fully embrace diversity of outcomes, expected and accepted a range of achievement and did not perceive this to compromise standards in their discipline, although one or two in AHSS were actively concerned about this variation above the threshold. Most were confident that, at least in their own field, standards were maintained by a common understanding of the threshold. Second, the flexibility of institutional regulations, while there is a strong argument for greater conformity, does currently give examiners flexibility to adjust their recommendations to

accommodate individual candidates' circumstances, including adjustment of 'minor corrections', which may be expanded or contracted depending on the examiners' approach and their degree of empathy with the candidate's circumstances. Greater consistency of time periods and the extent of amendments permitted in the minor corrections category would address inconsistencies that, especially when differentiating between those with different career aspirations – academic or non-academic – potentially suggests variation in standards and different forms of 'doctorateness'. The low incidence of failure is unsurprising, given progress hurdles now in place for all candidates, yet the underused award of a lower degree, perceived not to be a 'real' option by some examiners, does present arguments for considering it not as a 'consolation prize' but for some, a more positive outcome enabling them to progress. The acknowledgement by some examiners that they were aware some candidates were 'marginal' and perhaps should not have passed, is a concern and would be addressed by more positive use of the lower degree. Third, this chapter above all illustrates the many challenges inherent in differentiating the range of candidate achievement above the pass/fail threshold and the myriad of factors affecting each individual's progress towards completion. While there was almost no enthusiasm for grading candidates who pass, there was more support for rewarding exceptional candidates through some form of distinction recorded on the degree certificate and restricting the judgement to the thesis quality, rather than taking account of performance across the entire examination.

Chapter 8: Discussion of findings and critical reflection

This study took place at a time of change for the doctorate internationally, resulting in questions concerning the purposes and value of the PhD and its examination. Several factors combined to create challenges for the final examination, which remains largely unchanged since its introduction in the UK. These include a significant expansion in the number of PhD candidates and graduates and much greater diversity in their backgrounds and motivation, many seeking to differentiate themselves in the job market. Simultaneously, driven by economic considerations, government and related agencies have taken a more interventionist approach to doctoral education, aiming to emphasise so-called transferable skills related to employment. The introduction of more structured training aimed to satisfy employer expectations by developing these skills and led to the now widespread concept of the 'structured PhD'. This initiative coincided with major developments in the field of assessment in higher education and in particular the introduction of 'learning objectives' at all levels.

At the same time, low level though continuous criticism of the PhD examination process continues. Is it fit for purpose? Is it fair, reliable and valid? What is the purpose of the viva, and is it essential? With increased internationalisation comes an increase in comparisons with other systems of PhD examination (e.g. the 'European' model of public defence or the Australian model of examination of thesis only). These questions, and the fact that there has been no previous empirical study of the examination process, made the timing of the present study particularly relevant.

My main research question: ***'How does the PhD examination enable examiners to make judgements about the candidate and their work and is it still fit for purpose in the modern context?'*** allowed me to focus on the challenges faced by examiners, given the different contexts in which the PhD examination occurs, as shown by the case summaries in Appendix 1. I was particularly interested in how the thesis and viva contribute to the process and to explore why, even though the PhD involves structured training designed to deepen and increase consistency of professional and personal attributes, these qualities are under-emphasised in the examination. In this discussion I focus first on what I consider to be a holistic process that I have termed the 'thesis-viva nexus'. I then consider the 'attributes' sought by examiners in the thesis, viva, or both. Finally I explore examination outcomes, the

sustainability of a pass/fail threshold judgement and whether variations in quality (thesis and candidate) might be differentiated in the final result.

The PhD examination as a holistic process

Few authors, if any, have analysed the examination process holistically. Those who have published on the process, for example, Jackson and Tinkler (2004), or Denicolo (2003) often include it as part of a comprehensive exploration of many facets of the examination. My first subsidiary question – *In the examination process, what role is played by the thesis, what are the purposes of the viva, and what is the relationship between them?* – led me to explore the process as a temporal and conceptual ‘continuum of judgement’ (Figure 5.1, p.106). Strikingly, this established the strength of the thesis:viva relationship – the thesis-viva nexus. According to almost all my respondents, the thesis incontrovertibly dominates examiners’ judgements. This supports Denicolo’s assertion that most examiners and supervisors considered the thesis to form ‘*the major source of evidence for the quality of the research and the level of scholarship*’ (2003;84 and see Lovitts, 2007). The thesis-viva nexus strengthens the process by allowing examiners to reflect, confirm or review initial impressions and seek clarification from the candidate, leading to a rounded judgement.

My findings suggest the thesis-viva nexus is a key feature of the PhD examination in the UK, as described by this examiner, whose comments epitomise many of my findings, especially the impact of the viva:

The thesis describes the work done and [the candidate’s] written understanding. The viva tells you whether they’ve done it themselves and how well they’ve understood it. They have to meet basic criteria on both sides. But [the viva] might increase or diminish confidence in [the candidate’s] achievement. You are looking for an overall level of achievement and understanding, within the context of the thesis or viva. If they haven’t managed to get this across in the thesis, you’re looking for it in the viva. There is a basic level for both, as long as they can show enough understanding on one side or the other. Thesis plus viva is a good format. Having the viva is part of the student’s training – you are always trying to improve the student’s learning rather than judge them and the viva contributes to that. (BIO-Ex2)

Conceptually, the UK examination enables an integrated assessment of each candidate’s research, professional and personal attributes, as illustrated in Figure 8.1, the arrow representing an increasing emphasis on professional and personal skills as the continuum progresses.

Thesis	Viva
Higher quality thesis means more emphasis on peer-to-peer discussion of research results in the viva because of examiners' confidence in the candidate.	Viva becomes more important if the examiners have concerns after reading the thesis or require more information.
Much though not all evidence of research achievement is derived from the thesis, in addition to some of the candidate's professional skills, e.g. written communication.	Viva enables examiners to explore the candidate's research in more detail, check authorship and evaluate professional and personal attributes.

Figure 8.1: The thesis-viva nexus

I refer to this as a holistic approach to the examination, which was shown by my study to have an important role in assessing wide-ranging candidate attributes. The holistic nature of the judgement contributed to demonstrating validity, integrity and rigour in each of the ten vivas. Perhaps most significantly for standards, the viva assumes greater importance if the thesis is weak. This aligns with Jackson and Tinkler's (2001) findings, 29% of whose respondents saw it as an opportunity to provide formative guidance to candidates to strengthen their theses. In this study I observed that in the case of high quality theses, the viva took the form of a peer-to-peer discussion relatively early in the process, whereas if the thesis raised concerns, formative feedback increased, with examiners focused on corrections to improve thesis quality (Mullins and Kiley, 2002; Holbrook et al, 2004a, 2004b; Holbrook et al, 2014). Observations confirmed that formative feedback with an emphasis on improving the candidate's learning was a feature of all ten vivas, with the examination delivering both summative and formative assessment.

No clear consensus exists in the literature on whether the viva changes examination outcomes, as suggested by Berelson (1960) whose survey of graduate deans revealed a polarisation of views on this question. This polarisation is reflected in more recent studies. On one side, Noble (1994) suggests the viva is redundant, a view shared by the majority of Jackson and Tinkler's (2001) respondents, who suggested it rarely changed outcomes. In their study, 6% judged the viva to be completely unnecessary. In broad agreement, Lovat et al. (2015) contend that the

viva is rarely thought to contribute significantly to examiners' judgements. On the other hand, all acknowledge it is valuable for checking thesis authenticity, and authors who address the practicalities of the examination, such as Tinkler and Jackson (2004) and Phillips and Pugh (2010), show the viva has multiple purposes.

Most of my respondents agreed on the necessity of the viva, although as Lovat et al (2015) and Kyvik (2014) found, examiners tend to favour the system they are most familiar with.

Interestingly, candidates suggested that, were it not for the viva, their examiners would have had an incomplete understanding of their research and even Lovitts (2007), whose study focused almost entirely on the dissertation (thesis), thought it important for examiners to meet candidates. Significantly in this study, I observed that the viva could influence outcomes. Some did not go as expected and there were step changes in recommendations in at least two cases. I will make brief reference to three particular cases that show the viva as a dynamic event that makes a significant contribution to the examination. In case F, the outstanding quality of the candidate's film-making was reflected in their performance in the viva and although the thesis required more emphasis on the wider field of literature, during the viva the candidate demonstrated they possessed the knowledge missing from the text. In case G, on the other hand, the examiners came to the viva with concerns which, had they been dispelled by the candidate in the viva, would have resulted in minor corrections, but the candidate's responses served to demonstrate that more work was required and major corrections were recommended within a nine month period, surprising the supervisor but appreciated by the candidate. In case J, the candidate's thesis did not do justice to the methods employed in the study, nor to the results, which were considered by both examiners to be significant. The candidate had experienced disruption in the original supervisory arrangements and had no-one to ask when they had problems conducting experiments. The viva enabled the examiners to recommend major corrections to be completed within six months rather than recommending referral or re-submission, because they were reassured that the weaknesses in the thesis emanated from the lack of support the candidate experienced.

As one examiner remarked: *'I very much see the viva as a moderation process'* (COMP-EX1). This seems to me to be an important and significant role in the holistic process. In all cases I observed, the viva provided evidence of examiners moderating their thesis-based judgements, often disclosing new insights and providing an alternative context for candidates more proficient at oral than written communication to display the necessary attributes (Noble,

1994). An additional feature of the viva that emerged from the observations was as a learning opportunity for examiners, who in some cases learn more about the field from the candidate's research. This suggests the viva may be valuable for the academic enterprise and communicating research in the discipline, raising awareness of new knowledge and understanding in the field.

One of the criticisms of the current UK examination process is that arranging and funding the viva is increasingly expensive and time-consuming. Is it worth it? Does it make a difference? Do the outcomes justify the expense? In this study my respondents agreed on the moderating role of the viva and its important contribution to examination outcomes. Nevertheless in the UK, the viva process, described by one examiner as one of the most unregulated systems in higher education, does not fully meet current expectations of assessment. In particular, concerns remain regarding its closed and private nature. For this reason the past 10 or 15 years have seen the growing use of an independent chair or convenor, whose role is to mediate the conduct of the viva, a practice strongly encouraged by QAA (QAA, 2012a).

Convenors were present in half of the cases in this study; three were interviewed. From the observations I noted that convenors had a range of roles. These included introducing the examiners to the candidate; explaining they were present to ensure the smooth running of the examination and that the university's regulations were followed; informing candidates they could ask for clarification of examiners' questions, or for a break; advising examiners on protocol regarding citing and clarifying authorship of papers included as a thesis chapter; and advising candidates the viva was meant to be helpful rather than confrontational and that they would be asked to leave the room at the end while the examiners finalised their judgements. Most importantly, none of the convenors contributed to the judgement. One set out five clear purposes for their role:

To assure fair realisation of policy, a common framework for conducting the viva that is fair for all, irrespective of who the examiners are; that the candidate is comfortable and able to give of their best – is not distressed...to be aware of any medical condition a candidate might have and [make] allowances...to be watchful of changes in behaviour and aware of the candidate's needs. (AHSS-IC1)

The study helped to define the facilitative role of the convenor in real viva situations and demonstrate their importance in interpreting regulations, generally supporting both the examiners and the candidate and, crucially, assuring some consistency in the process.

Attributes sought by examiners

My second subsidiary research question addressed the question of attributes sought by examiners in the PhD examination – *What attributes do examiners seek in the thesis and the candidate and how are they demonstrated?* Given the initiatives on fostering so-called transferable or employment related skills (e.g. arising from the Roberts report), the developments in assessment concerning intended student learning outcomes (ILOs) and the doctoral qualification descriptor (QAA, 2008), I was interested to explore the attributes sought by examiners, in both the thesis and, in particular, the viva.

Despite the lack of evidence of examiners' adherence in the examination to doctoral ILOs, data relating to this research question generated a critical and positive finding – that examiners in all fields seek similar attributes in doctoral candidates. In the ten cases contributing to the study, this and my respondents' assertions that they rarely disagreed with their co-examiners, ameliorated concerns about a lack of formal adherence to doctoral ILOs. While the nature of each subject influences how examiners in different subjects go about testing for research and other attributes (Mullins and Kiley, 2002), as Chapter 6 shows, those contributing to this study exemplified much common ground in their approaches to evaluating 'doctorateness' (Denicolo and Park, 2010). Table 6.2 (p.188) presents four groups of core attributes (A1–A4) my respondents considered central to judgements. It aims to demonstrate two forms of integration in examiners' judgements. First, it describes the interdependent relationship between the thesis and the viva in the way judgements are made in all but the publishability category. One examiner described this relationship as: *'There are two sides to the candidate, aren't there? There's the writing and then there's the viva...those two are intimately linked...'* (AHSS-I2). Second, the table suggests that both research and professional and personal attributes contribute to judgements in each of the four categories, showing another facet of integration. Both forms of integration are important, but the second in particular suggests that 'doctorateness' is evidenced in both research and professional and personal attributes. For example, interviews and observations both indicated that professional and personal attributes such as creativity, flair, initiative and breadth, as well as technical ability, contributed to judging research competence. It was interesting to note that differences of interpretation or approach between AHSS and STEM respondents mainly concerned research attributes, whereas in the group of attributes relating to intellectual rigour, there were no obvious disciplinary differences.

Figure 8.2 describes how the boundaries between the thesis and viva and the attributes they evidence are blurred, supporting an integrated, holistic judgement.

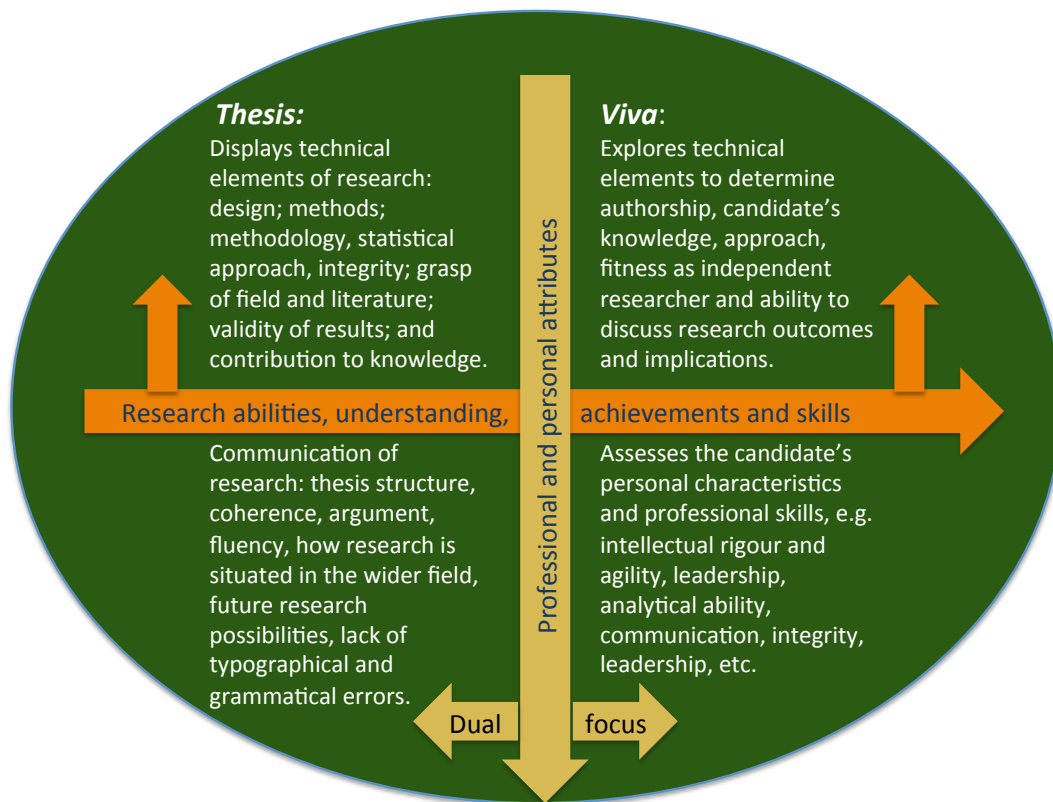


Figure 8.2: Integrated judgement

The study also contributes to the understanding and interpretation of the attribute of 'originality' in the context of the PhD, suggesting that more explicit definitions may be needed to differentiate between most candidates and those who demonstrate exceptional achievement (Poole, 2015; Johnston, 1997). The broadly opposite interpretations of 'originality' and 'contribution to knowledge' that exist among AHSS and STEM respondents suggest the need to include the subject context in any definitions. It is obvious, yet worth mentioning, that the nature of the candidate's subject will lead to development of different attributes, for example, excellent written and spoken communication skills in a modern languages PhD, high levels of numeracy in a mathematics doctorate or special appreciation of clinical considerations in healthcare practitioners. Some such attributes may be considered 'transferable' in the context of future employment. The initiative taken by the IUBMB for biological sciences (Chapter 2) exemplifies the advantages of introducing subject-specific ILOs in doctoral assessment, including the potential for increasing consistency and maintaining standards. Initiatives of this kind can only be successful, however, if they emerge organically

from subject communities. The IUBMB's Standards, encompassing research, professional and personal attributes, developed from concerns about the extent of diversity among the academic community (Lovitts, 2007) and in candidates embarking on PhDs in biology and the associated impact on standards and outcomes in the international higher education environment.

Examination outcomes: how do they reflect candidate achievement?

While examination outcomes initially reflect the diversity of achievement among PhD candidates through the range of recommendations open to examiners, the final result is pass, fail, or the award of a lower degree. The potential for differentiated outcomes is explored through the third research question: *How do different examination outcomes reflect candidates' achievements and is a pass/fail judgement still appropriate?* Variable achievement is unsurprising in a degree with such individuality and diverse entrants, yet there is a marked difference in standards above the threshold.

Despite the acknowledged differences among candidates, and the use of classification systems for most other degrees in higher education, I initially found it surprising that there was little support for grading the PhD. Most respondents considered the threshold judgement unproblematic and one or two even saw multiple levels of attainment as an inevitable and positive element of the examination given the individuality of the PhD, though some would prefer a narrower range of achievement. The pass/fail threshold is considered to maintain an absolute standard, once thesis corrections have been completed.

Closely monitoring interactions in the viva enabled me to gauge that, while all the theses involved in the study were considered 'passable' once corrections had been completed, and allowing for inter-subject differences, there was some variation in: volume and quality of research data; how well results were analysed and communicated; the extent of candidates' knowledge of the wider field; and the composition and structure of the thesis. This was unsurprising given disciplinary conventions, candidate individuality and the uncertain nature of research. Several of my respondents noted that outcomes are affected by the uniqueness of each candidate's experience, which was evident even from the small but varied sample of ten cases. Others contended that comparability of attainment exists within and across disciplines, partly because candidates have researched a topic in a structured training environment for 3-4 years, passing progress hurdles along the way, and also because of the moderating influence

of corrections on thesis standards. These were convincing arguments for retaining a threshold judgement, especially when combined with my observations. As we have seen, however, recommendations for corrections may be inconsistent either because of regulatory variation, disciplinary conventions or examiner approaches.

Borderline theses were of particular interest, given the greater reliance on the viva in such cases. The few marginal cases mentioned by respondents appeared to demonstrate a lack of breadth or depth and frequently involved referral and a second viva. It is unclear how much thesis standards are currently affected by pressures for completion and could be improved if more time were available, a question raised by only one of my respondents, although one candidate mentioned they had struggled to finish in the time available as their sponsor had only provided funding for three years (Case D).

Lovitts' (2007) respondents confirmed they hardly ever failed a dissertation, as did examiners in this study, especially in cases where unsatisfactory supervision results in an inadequate thesis. Respondents' positive perspectives on achievement emerging from this study are supported by the very few referrals, lower degrees or failures they had experienced in candidates they had examined or supervised (Tables 7.2 and 7.3, pp 204 and 209). Only one (STEM) examiner alluded to the possibility of failure for lack of originality, which they had experienced but which was rare (Holbrook et al, 2004a). Most candidates pass with minor corrections. This was true of the candidates I observed, with eight out of ten receiving this recommendation.

On one hand, the low number of referrals, lower degrees and failures is reassuring for standards yet, even if rarely, my respondents had experience of candidates who should have failed, but were referred and subsequently passed having amended their thesis using very detailed formative feedback from their examiners. While this brings the thesis to the necessary standard to pass, is it fair to other candidates who do not receive extensive support, and should more use be made of lower degrees? Examiners are well aware of the impact of failure on the candidate who has spent many years pursuing their goal only to be denied a PhD at the final hurdle, and often face moral dilemmas in such situations – should they rigorously maintain the standard or support the candidate to produce a passable thesis?

Given the acknowledged variability in thesis quality, the lack of support for grading or distinctions among the majority of examiners, supervisors and candidates was striking, yet it

was explained by their counter arguments. The challenge of creating criteria that could apply equally across disciplines, the tension between achieving a high quality thesis within the time available and focusing on publications during the programme, the potential for grade inflation and the difficulty of advising candidates who enquired how to raise the standard of their work to distinction level, all challenge the concept of grading. The 'licence to practise' comparison with other threshold-based qualifications such as medical degrees is another reason for retaining a pass/fail outcome. Candidates' arguments for retaining the threshold judgement were also convincing and included the already high academic status of the PhD, the variability of candidate experiences, and the straightforwardness of the current system as they saw it.

One of the most interesting aspects of exploring examination outcomes was to evaluate if each individual's 'doctorateness' was established and evidenced in a way that could be interpreted by any stakeholder, including employers. While examiners demonstrated that they were assessing the candidate as an individual, their overt focus in the viva was on the quality of the research and the contribution to knowledge. This tension is sometimes referred to as 'output' versus 'outcome' or 'product' versus 'process', some suggesting that the focus of assessment should be on the individual, or outcome, rather than the product (the thesis and/or publications). Should the contribution of professional and personal attributes be more overtly recognised in the judgement and in examiners' reports?

Challenges and limitations

I believe that mine is the first study involving viva observations by an 'outsider' – an independent observer who, until embarking on the research, had no practical experience of vivas either as a candidate or examiner. Yet my position was also a form of 'insider-outsider' researcher. On one hand, as an independent researcher I was an 'outsider'. Conversely, two features of my situation constituted 'insiderness': first, my professional role in the field of doctoral education and second, my position of doctoral candidate undergoing the examination. This dual identity created both challenges and unique insights.

In Chapter 4 I highlighted two potential weaknesses in my research design: that I was likely only to be granted access to vivas where everyone was confident the candidate would pass and therefore relied on opportunistic sampling; and that behaviour during the observations was likely to be changed by my presence. The first was an inevitable flaw in my methodology given the sensitivity of the topic, and potentially more significant than the second, since

examiners and supervisors would not agree to my observation if they had any concerns about the outcome of the examination. The inbuilt uncertainty of PhD research that inevitably affects the examination, however, meant none of the examiners and supervisors could be certain of the results, even if they were expected to be positive. Evidence for this uncertainty is apparent from two candidates in particular. The first, Case G, received major corrections rather than minor as a result of what the examiners discovered during the viva, while the second, Case J, was always expected to receive at least major corrections based on thesis quality. In Case J, the examiners were much impressed with the candidate during the viva: they nevertheless recommended major corrections to assure the standard of the thesis, yet as a result of the viva knew the candidate was more than worthy of the PhD. These and other cases convinced me that my research design was sufficiently robust to deliver authentic, but not generalisable, results. The second design flaw was that it was impossible to know if my presence in the vivas would alter the actors' behaviours. This is an unavoidable problem with all observation studies, though potentially less so in the five cases where a convenor was present. However, in all the vivas, once they were underway, participants appeared unaware of my presence.

These potential weaknesses were compounded by the access constraints I experienced and in the earlier stages I doubted the approach would produce sufficient relevant data. The difficulties of access, described in Chapter 4, were hugely challenging, both for me as a researcher and for the study. Nevertheless, the risks inherent in my research design proved worth taking for the almost unprecedented access to viva observations. Even though I appreciated theoretically the privacy and confidentiality involved in viva examinations, I had not anticipated taking more than two years to finish my field work. Having completed the two pilot observations in November and December 2011, it was almost a year before I was able to embark on two more. There was then a gap of three months before I completed two further observations in March and April 2013, followed by two more in June and July the same year. Another gap of nine months elapsed before I completed the final two observations. These time periods bring into focus the periods of negotiation undertaken to achieve the observations, but do not express the frustration and disappointment involved in attempting to conduct them. During the study I realised it had been ambitious to plan to observe fifteen vivas, so when I managed to secure ten with a reasonable amount of disciplinary variation it was both satisfying and a relief to decide I had generated sufficient data.

The observations themselves were fascinating and brought the examination to life. Some were physically demanding, lasting more than three hours, during which I could not move or make any noise, while taking detailed notes of the discussion. The note taking was particularly difficult in subjects that I knew nothing about, for example, immunology, but I was able to make sense of my notes afterwards and the subject matter did not prevent me from identifying conceptual similarities and differences in the examinations. Importantly for the methodology, the observations enabled me to gain a deeper understanding of how the viva contributes to examiners' judgements, to probe my respondents more intelligently in interviews and compare my impressions with theirs, and to appreciate candidates' experiences of the 'rite of passage'.

Conclusion

One of the most surprising outcomes of my research, given that the PhD examination contrasts with almost all other forms of assessment in higher education, was that examiners in particular did not highlight its uniqueness or criticise the examination, suggesting they accept the process and outcomes as they are. Even those who thought the examination could be improved mentioned small changes, such as an additional examiner or an assessment panel rather than two examiners, and did not compare it with other forms of assessment. Was this because they were confident the examination enabled them to make accurate and defensible judgements, or because the degree is so uniquely rooted in research that they consider a completely different form of assessment is appropriate, or both? Examiners believe that common standards exist and are related to expectations in the field, and the IUBMB Standards suggest it is possible to define research, professional and personal attributes that can be agreed upon by the international academic community.

My findings indicate that some of the challenges explored earlier in this chapter are balanced by: the confidence of examiners that standards in their discipline are almost always upheld; their obvious wish to support candidates to succeed combined with rigorous questioning in the examination; and the formative feedback that influences the quality of all theses. Nevertheless, given that numbers and diversity of PhD candidates have increased, with the majority seeking non-academic careers, it is even more important that the attributes displayed by candidates in the thesis and the viva are made more transparent. The doctoral examination and its outcomes need to maintain credibility, not just within academia but externally, so that all stakeholders understand the concept of 'doctorateness' and how it is judged.

Chapter 9: Conclusions

The timing of this study, which explored the process and outcomes of the PhD examination and the candidate attributes sought, coincided with significant changes in doctoral education and continuing low level concerns about the nature of the examination, particularly the viva. I framed my study around the research question: *How does the PhD examination enable examiners to make judgements about the candidate and their work and is it still fit for purpose in the modern context?* The wider question of fitness for purpose pertains to the lack of change in the examination for over 50 years juxtaposed with increased focus on graduate transferable skills. The study enabled me to look in detail at ten cases as an observer of the viva examination and to interview over 40 actors involved in the process (as examiners, supervisors, candidates, or convenors). The amount, quality and relevance of data generated by the study enabled me to shed light on how examiners judge candidates' achievements. Data also demonstrated the strengths of the examination and some of its continuing weaknesses. The findings chapters provide three perspectives on the research question, showing respectively how the process facilitates judgements, the attributes on which judgements are based, and to what extent outcomes can be viewed as fair and possessing academic integrity, simultaneously reflecting the purposes of the modern PhD.

The PhD examination process

The UK PhD examination consists of a two-part process, with evaluation of the thesis frequently separated from the viva examination by several weeks if not months. This may be considered to form a continuum of judgement. My study supports the idea of the examination as a holistic process, which I refer to as the 'thesis-viva nexus'. Participants were in no doubt as to the importance of the viva, which fulfils several different purposes, a core function being to enable examiners to finally moderate their judgements. Moderation occurs over time: after the two examiners have read the thesis independently and formed individual judgements, which they may adjust as a result of re-reading, they meet to discuss their individual judgements. The meeting may lead to further adjustments, from which a joint preliminary judgement emerges. The viva enables additional and crucial moderation as examiners' judgements are influenced by what they learn in the viva. The final stage of moderation occurs during their post-viva discussion and joint report. In addition to establishing the viva's moderating role, the study demonstrated how it enables examiners to authenticate the thesis, and assess some of the candidate's professional and personal attributes. The viva also provides

candidates with an occasion to mark their significant transition to 'doctor' and to engage in an 'expert' discussion of their research.

Attributes contributing to 'doctorateness'

In making judgements about candidates, examiners are asking what attributes PhD graduates might be expected to have on completion of the doctorate. This question has become more important with the introduction of doctoral qualification descriptors by QAA and equivalent organisations internationally. It is also significant in the light of developments in assessment practice that focus on intended learning outcomes (ILOs). A striking finding of the study was the extent of agreement among respondents on the key attributes sought in PhD candidates and what signifies an exceptional candidate.

The essence of the PhD remains its uniqueness in training individuals to become independent researchers, while also preparing graduates for many other careers, so it is unsurprising yet essential that the attribute of originality and/or a contribution to knowledge remains the principal criterion, especially for candidates. However, the study suggested that the publishability of the work, linked with a range of other professional and personal attributes, may be equally important in the overall assessment. Crucially, the findings suggest that both the individual and their work are being assessed equally during the examination (Table 6.2; Figure 8.1; Lovitts, 2007; Hall, 2006; Denicolo, 2003; Mullins and Kiley, 2002), even though this dual purpose of the process has not yet been fully articulated and the challenge of communicating the essential attributes of 'doctorateness' remains. However, the study provides evidence that examiners' judgements span research, professional and personal attributes (Table 6.2; Figure 8.1). It suggests a need for greater emphasis on the characteristics that signify 'doctorateness' and how they combine to form an integrated judgement. Given the significant developments in and funding allocated to programmes supporting the acquisition of 'transferable' or 'employment related' skills in the UK, it is pertinent to ask if and how these might be assessed and in particular where they might feature in the final PhD examination. If expectations are that PhD graduates should possess certain skills including and additional to those prescribed in the doctoral qualification descriptor (QAA, 2008), should these be articulated and evidenced in the final examination? This remains a relevant and important question.

Examiners' judgements of PhD outcomes

The UK continues to use a pass/fail outcome for the PhD, implying that examiners are required to have a clear notion of the 'threshold' standard and by implication an understanding of what constitutes doctoral level in both thesis and viva. There is, however, a general recognition of variability in candidate achievement. This recognition of differential attainment was demonstrated in the study by examiners' use of broad terms such as: exceptional, 'middle of the road' and borderline, reflecting variable achievement across disciplines. This is not to suggest, however, that standards are compromised, especially since initial judgements based on the candidate's work are mediated through formative feedback in the viva and in examiners' recommendations for corrections, which influence the final thesis. Nevertheless, the tension between achieving fair outcomes while maintaining standards is evident from my findings, in particular, use of the options open to examiners at the end of the viva, with inconsistency of institutional guidance raising questions of differential treatment of candidates, especially in the interpretation of minor corrections.

Despite the acknowledged variability in candidate achievement, my study suggested that the appetite for grading the PhD is limited. Arguments against grading were strong and indicated the belief that a threshold judgement remains appropriate, especially since for many PhD graduates, the degree is conferred at the beginning of their professional life, before their principal achievements, even though it has life- and career-enhancing implications in any context. My respondents, while discounting grading, did not, however, rule out the possibility of rewarding exceptional achievement, for example through starred degrees or distinctions, even though other options are open to high achieving candidates, such as supervisor or examiner references, or mentioning on a CV having passed with no corrections. Against distinctions for the minority was the concern about devaluing others' achievements when all have been awarded a degree with high academic status. Nevertheless, examiners' ability to recognise exceptional achievement and the different interpretations of originality or a contribution to knowledge that clearly differentiate between high achieving and all other candidates that is shown in Chapter 6, suggests that some recognition such as a distinction could reward exceptionality in research achievement, as demonstrated by thesis quality.

The potential impact of the study on the PhD examination

While the study does not suggest radical changes to the examination format, it highlights possible strategies for assuring the consistency of the viva and particularly examiner approaches, that could increase confidence in the process and its outcomes. The use of an independent chair or convenor for PhD vivas has been clearly recommended in QAA guidance. In five of the ten cases in this study, the presence of a convenor, intended to help assure the appropriate conduct of vivas, adherence to institutional policy and address any inconsistencies in examiners' approaches, appeared to work well. Convenors in this study also had an important role in assuring the candidate's comfort and understanding of the process which otherwise are the responsibility of the examiners. An alternative for institutions for which the concept of a convenor is problematic would be to record all vivas. Recordings can be accessed when outcomes are not straightforward or concerns arise, or for random quality assurance checks. Recording, however, removes the option for the independent chair to facilitate, be consulted or take action during the viva should the need arise.

Having undertaken this small study, my overall conclusion is that the PhD examination in its current form did enable examiners to make fair and rigorous judgements in the ten cases. However, two further measures could improve consistency in how the viva is conducted. The first, to introduce compulsory training for doctoral examiners, could increase consistent understanding of the purpose and process of the viva and its relationship to the thesis. However, such training could make unrealistic resource demands and compromise institutions' ability to source examiners. The second is to invite all candidates to briefly introduce within a fixed time, for example, ten minutes, the most important elements of their research at the beginning, speeding up the process and increasing the evidence available to assess attributes, particularly some professional and personal skills.

Two further changes in particular might increase transparency of the examination and its outcomes for all stakeholders, including graduates and employers. The first could be achieved by augmenting ILOs in national and institutional doctoral assessment criteria, emphasising the interdependence and equivalence of attributes sought by examiners, without in any way detracting from the core requirement of originality or a contribution to knowledge. A second step could be to require examiners to identify explicitly research, professional and personal attributes in their reports and where they are evidenced, reflecting the holistic approach apparent in the ten cases that suggested some professional and personal attributes are

important for research competence. Approaches such as the IUBMB's doctoral standards statement have potential to increase disciplinary consistency, while any broadening of ILOs could incorporate alternative definitions of originality or a contribution to knowledge, to help clarify disciplinary and other interpretations, for both academics and candidates. These measures would focus doctoral outcomes more broadly, accentuating their relevance to all careers and employers, academic and non-academic. From examiners' reports, candidates could provide evidence of 'doctorateness' in research-specific, professional and personal qualities, to supplement other evidence acquired during their programme.

The study also suggested that a UK-wide point of reference on the definition and interpretation of minor and major corrections, to include examples of when they are appropriate with indicative time periods, could serve a useful purpose in encouraging greater consistency and moderating recommendations. However, this could restrict the flexibility currently available to examiners that enables them to reflect individual candidates' circumstances and support them to complete corrections.

Future research

The scarcity of empirical research on the doctoral examination suggests several possible areas and approaches for future research. Mine was a small-scale study involving ten cases (and their associated examiners, supervisors and candidates). It would be important to discover how far these findings apply to a wider group. Important questions concern the role and purpose of the viva; the assessment of professional and personal attributes; the qualities that represent 'doctorateness' and whether they vary among disciplines; and the feasibility of introducing distinctions for high achieving candidates.

Questions specific to each of these areas include: the possibility of extending quality assurance measures such as the universal introduction of viva convenors, recording vivas, or more widespread examiner training; creating enhanced national ILOs for doctorates that would emphasise the interdependence of research, professional and personal attributes; encouraging doctoral examiners to make explicit reference to professional and personal attributes in their reports; introducing some form of distinction for exceptional research achievement in research as evidenced by thesis quality and linked to subject-specific definitions of originality and a contribution to knowledge; exploring the feasibility of introducing discipline-specific doctoral attributes developed by professional bodies or learned societies; and whether it would be

advantageous to introduce guidelines to encourage greater consistency in the use of minor and major corrections while maintaining enough flexibility to accommodate candidates' individual circumstances.

It would also be of interest in the context of the final examination to explore whether the changing nature of research, for example, the increasingly interdisciplinary nature of enquiry, should have a bearing on the process, including the format of the thesis and acknowledging multiple contributions.

The limitations of the study, such as the challenges involved in viva observations, particularly those outlined above, are likely to affect any future research based on similar methods, yet the insight available from independent observation would be difficult to replicate using other forms of enquiry. Nevertheless, obtaining data through questionnaires, for example, would overcome the problems of access experienced in this study. Had I followed my original plan to use questionnaires, they could have generated potentially generalisable results through surveying larger numbers of examiners, supervisors and candidates, with particularly interesting responses followed up through interviews. Any future research could combine observations, questionnaires and interviews to triangulate results.

Conclusion

The strengths of the UK process were shown by the study to be: the principal requirement for the candidate to contribute to knowledge in the field, as is appropriate for the only degree that confers the status of being able to conduct research independently; the involvement of at least two independent examiners who must agree on the outcomes; the time allowed for reflection and examiner discussion to synchronise judgments during the continuum; the opportunity for examiners to assess candidates face to face in the viva, enabling moderation of initial judgements (Kyvik, 2014), and authentication of the thesis; the contribution of examiner feedback to improving the standard of the final thesis; and providing candidates with a 'rite of passage' after years of work. Interview data suggest that, additionally, progress examinations during the programme help to assure the quality of the candidate's intellectual engagement and ability to produce a coherent thesis. Another positive feature is that the UK system does not permit examiners to disclose their recommendations to candidates until the end of the viva, when they are as fully informed as possible about the candidate and their research. At this point, the thesis could be considered 'unfinished', with formative feedback during the

viva, whether relating to high quality or borderline theses, informing corrections recommended in examiners' reports that contribute to maintaining standards, especially in borderline theses, and introducing greater consistency in the work of candidates who pass.

On the other hand, the final examination was described as an unregulated, '*crude system*' (AHSS-E1), which could be improved. Many might agree with this description of what is seen as an unreliable process for such a high stakes assessment, even though the uniqueness of each degree and the individuality of each candidate's training and supervision render it challenging to improve the process.

The body of research on the PhD examination prior to my study appears to have had only a minimal impact on practice. This is unsurprising given that the examination is a complex social phenomenon that is heavily influenced by the characteristics of the PhD itself, as Morley identified:

The imperative for originality in doctoral studies and the need for specialist examiners means that judgements of worth are difficult to standardise within a framework of measurement and outcomes-based approaches. (Morley et al., 2002:268)

The nature of the PhD, however, does not prevent changes to the examination and its outcomes that respect its individuality as a qualification and the imperative to train independent researchers while also demonstrating how graduates acquire numerous transferable attributes. The study did not raise major concerns about standards and displayed the many strengths of the examination, including the effectiveness of the reflective process. Additionally it demonstrated a surprising degree of consistency in the wide-ranging attributes sought by examiners, many of them transferable to other contexts. The challenges to the credibility and integrity of examination outcomes could be addressed by some of the actions suggested above, for example, increasing quality assurance of vivas and consistency of outcomes, supported by further research. Importantly, the study shows that although '*doctorateness*' is conceptually a key element of the examination, it needs to be defined and articulated to ensure it is understood by and meaningful to the academic community and those outside it, especially employers. Adapting the UK PhD examination to enable full recognition of the breadth of candidates' achievements need not compromise its fundamental purpose as an evaluation of independent research capability.

References

Anonymous Academic (2016) PhD vivas are unfair and I am paying the price. *The Guardian*, 26 February 2016.

Atkinson, P and Delamont, S (1985) Bread and dreams or bread and circuses? A critique of 'case study' research in education. In M Shipman (Ed) *Educational Research* (pp 26-45). London, Falmer.

AQF (2013) *Australian Qualifications Framework* 2nd Edition [Online]
Available from: <http://www.aqf.edu.au/> [Accessed 12 January 2015]

Australian Universities Teaching Committee (2002) *A comparison of norm-referencing and criterion-referencing methods for determining student grades in higher education* [Online]
Available from: www.cshe.unimelb.edu.au/assessinglearning
[Accessed 08 January 2016]

Baldacchino, G (1995) Reflections on the status of a doctoral defence. *Journal of Graduate Education*, Winter 1995, 1:3, 71-76.

Bassey, M (1999) *Case Study Research in Educational Settings*. Philadelphia: Open University Press.

Bassnett, S (2014) Cavalier attitudes lead to uncivil practices in the conduct of vivas. *Times Higher Education*, Issue 2152, 15 May 2014.

Becher, T, Henkel, M and Kogan, M (1994) *Graduate Education in Britain*. Higher Education Policy Series 17. London: Jessica Kingsley Publishers Ltd.

Bennet and Turner (2013) *PRES 2013: Results from the Postgraduate Research Experience Survey*. York: Higher Education Academy

Berelson, B (1960) *Graduate education in the United States*. New York, London: McGraw-Hill, (pp 199-215).

Bhaskar, R (1978 and 2008) *A realist theory of science* (2nd edition). London: Harvester Wheatsheaf.

Bhaskar, R (1986) *Scientific realism and human emancipation*. London: Verso.

Bhaskar, R (1989) *The possibility of naturalism: a philosophical critique of the contemporary human sciences*. New York; London: Harvester Wheatsheaf.

Biggs, J (2003) *Teaching for Quality Learning at University: what the student does*. 2nd Edition. Maidenhead: Open University Press.

Biggs, J and Tang, C (2007) *Teaching for Quality Learning at University: what the student does*. 3rd ed. Maidenhead: Open University Press, McGraw Hill.

Birkbeck, University of London (2018) Criteria for the award of a PhD [Online] Available from: http://www.bbk.ac.uk/mybirkbeck/services/administration/assessment/award-criteria/phd_criteria

[Accessed 23 September 2018]

Blaikie, N (1993) *Approaches to social enquiry*. Cambridge: Polity.

Bloxham, S and Boyd, P (2007) *Developing effective assessment in higher education: a practical guide*. Maidenhead: Open University Press.

Bologna Working Group on Qualifications Frameworks (2005) *A Framework for Qualifications of the European Higher Education Area*. Copenhagen: Ministry of Science, Technology and Innovation. [Online] Available from:

http://ecahe.eu/w/images/7/76/A_Framework_for_Qualifications_for_the_European_Higher_Education_Area.pdf

[Accessed 23 September 2018]

Borkowski, N (2006) in Maki, P and Borkowski, N (Eds) *The Assessment of Doctoral Education: Emerging Criteria and New Models for Improving Outcomes*. Sterling, Virginia: Stylus. Introduction.

Boud and Falchikov (2006) Aligning assessment with long term learning. *Assessment and Evaluation in Higher Education*, 31: 4, 399–413.

Boud and Lee (2009) (Eds) *Changing Practices of Doctoral Education*. Oxon: Routledge.

Bourke, S and Holbrooke, A (2013) Examining PhD and research Masters theses. *Assessment and Evaluation in Higher Education*, 38: 4, 407–416.

BPS (2017) *Guidelines for Assessment of the PhD in Psychology and related disciplines* [Online] <https://www1.bps.org.uk/system/files/Public%20files/Policy/Assessment%20of%20PhD%20on%20WEB.pdf>

[Accessed 23 September 2018]

Broadfoot, P (2000) in Filer, A (Ed) *Assessment: Social Practice and Social Product*. London: Routledge Falmer.

Brown, G (1997) What is assessment? In *Assessing student learning in higher education*, Chapter 2. London: Routledge.

Brown, G, with Bull, J and Pendlebury, M (1997) *Assessing student learning in higher education*. London: Routledge.

Bryan, C and Clegg, K (eds.) (2006) *Innovative Assessment in Higher Education*. Abingdon: Routledge.

Bryman, A (2008). *Social Research Methods*. 3rd edition. Oxford: Oxford University Press

Burniston, S, Rodger, J and Brass, J (1999) *Enterprise in Higher Education – changing the mindset*. Research Brief No.117, York Consulting Limited [Online] Available from:

<http://webarchive.nationalarchives.gov.uk/20130323034021/https://www.education.gov.uk/publications/eOrderingDownload/RB117.pdf>

[Accessed 23 September 2018]

Carter, S (2008) Examining the doctoral thesis: a discussion. *Innovations in Education and Teaching International*, 45:4, 365-374.

CGS (2016) *Thought leaders convene to consider the future of the doctoral dissertation*.

Workshop Proceedings, US Council of Graduate Schools [Online] Available from:

<https://cgsnet.org/thought-leaders-convene-consider-future-doctoral-dissertation>

[Accessed 23 September 2018]

Clarke, G and Lunt, I (2014) The concept of 'originality' in the PhD: how is it interpreted by examiners? *Assessment & Evaluation in Higher Education*, 39: 7, 803-820.

Committee of Vice-Chancellors and Principals (1988) *The British PhD*. London: CVCP

Crotty, M (1998). *The Foundations of Social Research: Meaning and perspective in the research process*. St Leonards: Allen and Unwin.

Delamont, S., Parr, O. and Atkinson, P. (1998) Creating a delicate balance: the doctoral supervisor's dilemmas. *Teaching in Higher Education*, 3, 157-172.

Delamont, S, Atkinson, P, and Parry, O (2000) *The Doctoral Experience: Success and Failure in Graduate School*. New York: Falmer Press.

Denicolo, P, Fuller, M and Boulter, C (2000) Sharing reflections on the experience of doctoral assessment: the voices of supervisors and examiners. Paper presented at the British Educational Research Association (BERA) National Event, 2nd June, published in Fisher, J and Cornelius, N (Eds) *Challenging the Boundaries: Personal Construct Psychology Perspectives for the New Millennium*, EPCA Publications, Farnborough, pp.61-76.

Denicolo, P (2003) Assessing the PhD: a constructive view of criteria. *Quality Assurance in Education*, 11 (2), 84-91.

Denicolo, P and Park, C (2010) *Doctorateness – an elusive concept?* Gloucester: QAA

Dowling, P and Brown, A (2010) *Doing research/reading research: re-interrogating education*. 2nd ed. London, New York: Routledge.

EUA (2005) Salzburg Principles [Online] Available from:

<https://eua.eu/resources/publications/626:salzburg-2005---conclusions-and-recommendations.html>

[Accessed 23 September 2018]

EUA-CDE (2016) Doctoral Education – Taking Salzburg Forward: Implementation and new challenges [Online] Available from:

<https://eua-cde.org/component/attachments/attachments.html?id=53>

[Accessed 28 September 2018]

ESRC (1987) *The Social Science PhD. The ESRC Inquiry on Submission Rates* (Winfield Report). London: Economic and Social Research Council.

EUA-Council for Doctoral Education (2016) Doctoral Education – taking Salzburg forward: implementation and new challenges [Online] Available from:

<https://eua.eu/resources/publications/354:doctoral-education-taking-salzburg-forward-implementation-and-new-challenges.html>

[Accessed 23 September 2018]

Flyvbjerg, B (2006) Five misunderstandings about case-study research. In Atkinson, P and Delamont, S (Eds) *Sage Qualitative Research Methods*. Thousand Oaks: Sage Publications.

Funk, K and Klomparens, K (2006) Using the Assessment Process to Improve Doctoral Programs. In Maki, P. and Borkowski, N. (Eds) *The Assessment of Doctoral Education: Emerging Criteria and New Models for Improving Outcomes*, Sterling, Virginia: Stylus. Introduction.

Gibney, E (2013) Are PhD vivas still fit for purpose? *Times Higher Education*, Issue 2098, 25 April 2013.

Gomm, R, Hammersley, M, and Foster, P (2000) (Eds). *Case Study Method: Key Issues, Key Texts*. London: Sage.

Griffiths, A (2014) *The birth of the Quality Agencies*. Gloucester: Quality Assurance Agency for Higher Education.

Guba, E and Lincoln, Y (1994). Competing Paradigms in Qualitative Research, in N.K. Denzin and Y.S. Lincoln (Eds). *Handbook of Qualitative Research*. Thousand Oaks, California: Sage.

Guccione, K and Bryan, W (2017) To find out if a PhD is worth it, look beyond academia. *Research Fortnight*, 18 January 2017, p.22.

Hall, F (2006) Canadian practices related to the examination of PhD theses. *Report for Canadian Association of Graduate Studies*. Provided by Professor G Slater, Dean of the Graduate School, University of Ottawa, as a result of an informal interview 01 June 2011.

Hammersley, M. (2010) What is Qualitative Inquiry? Presentation, in: *National Centre for Research Methods Festival*, 5-8 July 2010, St Catherine's College, Oxford.

Hammersley, M. and Gomm, R. (2000) Introduction. In Gomm, R., Hammersley, M. and Foster, P. (Eds) *Case Study Method: Key Issues, Key Texts*. London: Sage.

Hartley, J and Jory, S (2000) Lifting the veil on the viva: The experiences of psychology candidates in the UK. *Psychology Teaching Review*, 9 (2), 76-90.

Hartley, J and Fox, C (2004) Assessing the mock viva: the experiences of British doctoral students. *Studies in Higher Education*, 29 (6), 727-738.

HEFCE (1996) *Review of Postgraduate Education*. HEFCE report M 14/96.

HEFCE (2003) *Improving standards in postgraduate research degree programmes: Formal consultation*. Bristol: HEFCE 2003/23.

HEFCE (2005) PhD research degrees: entry and completion (2005/02) [Online] Available from: http://webarchive.nationalarchives.gov.uk/20120118184719/http://www.hefce.ac.uk/pubs/hefce/2007/07_29/

[Accessed 23 September 2018]

HEFCE (2007a) PhD research degrees: update. Issues paper 2007/28 [Online] Available from: http://webarchive.nationalarchives.gov.uk/20120118175646/http://www.hefce.ac.uk/pubs/hefce/2007/07_28/

[Accessed 23 September 2018]

HEFCE (2007b) Research degree qualification rates. Issues paper 2007/29 [Online] Available from:

http://webarchive.nationalarchives.gov.uk/20091112192415/http://www.hefce.ac.uk/pubs/hefce/2007/07_29/

[Accessed 23 September 2018]

HEFCE (2013) Postgraduate education in England and Northern Ireland: Overview report, 2013. Issues paper, July 2013/14. [Online] Available from:

<http://webarchive.nationalarchives.gov.uk/20130704105924/http://www.hefce.ac.uk/pubs/>

[Accessed 23 September 2018]

HEFCE (2015) Higher Education in England: Key Facts, 2014/15. Figure 10: UK and other EU postgraduate research entrants by mode of study, 2002-03 to 2014-15. [Online] Available from:

<http://webarchive.nationalarchives.gov.uk/20180405122141/http://www.hefce.ac.uk/analysis/HEinEngland/HEENG/>

[Accessed 23 September 2018]

HEFCE (2016) National Student Survey results 2016. [Online] Available from:

<http://webarchive.nationalarchives.gov.uk/20170712122512/http://www.hefce.ac.uk/it/nss/results/>

[Accessed 23 September 2018]

Henwood, K (1996) *Qualitative Inquiry: perspectives, methods and psychology*. In Richardson, J *Handbook of Qualitative Research Methods for Psychology and the Social Sciences*. Leicester: British Psychological Society.

HESA (2016) Destinations of Leavers from Higher Education in the United Kingdom for the academic year 2014/15. Statistical First Release SFR237 [Online] Available from:

<https://www.hesa.ac.uk/news/30-06-2016/sfr237-destinations-of-leavers>

[Accessed 23 September 2018]

HESA (2017) Higher Education Statistics for the UK 2015/16. Table A: HE students by level of study, mode of study and HE provider type 2011/12 to 2015/16 [Online]

Available from: <https://www.hesa.ac.uk/data-and-analysis/publications/students-2015-16>

[Accessed 6 November 2017]

Higher Education Commission (2012) *Postgraduate education: an independent enquiry by the Higher Education Commission*. [Online] Available from:
<https://www.policyconnect.org.uk/hec/research/report-postgraduate-education>
[Accessed 23 September 2018]

Hodson, L and Buckley, A (2011) Postgraduate Research Experience Survey: 2011 Results. [Online] Available from:
https://www.heacademy.ac.uk/system/files/pres_report_2011_0.pdf
[Accessed 29 September 2018]

Holbrook, A and Bourke, S (2004) An investigation of PhD examination outcome in Australia using a mixed method approach. *Australian Journal of Educational and Developmental Psychology*, 4 (2004), 153-169.

Holbrook, A, Bourke, S, Lovat, T and Dally (2004a) Investigating PhD thesis examination reports. *International Journal of Educational Research*, 41 (2004) 98-120.

Holbrook, A, Bourke, S, Lovat, T and Dally (2004b) Qualities and Characteristics in the Written Reports of Doctoral Thesis Examiners. *Australian Journal of Educational and Developmental Psychology*, 4 (2004), 126-145.

Holbrook, A Bourke, S, Lovat, T and Fairbairn, H (2008) Consistency and inconsistency in PhD thesis examination. *Australian Journal of Education*, April 2008, 52:1, 36-48.

Holbrook, A Bourke, S, Fairbairn, H and Lovat, T, (2014) The focus and substance of formative comment provided by PhD examiners. *Studies in Higher Education*, 39: 6, 983-1,000.

House of Commons Innovation, Universities, Science and Skills Committee (2009): Students and universities. Eleventh report of the session 2008 09 HC 170-1. London, The Stationery Office Ltd.

In our time. Originality (2003) [podcast] BBC Radio 4. Broadcast 20 March. Available from:
<http://www.bbc.co.uk/programmes/p00548vy>
[Accessed 23 September 2018]

IUBMB (2011) *Standards for Doctoral Degrees in the Molecular Biosciences*. Recommendations of the Committee of the International Union of Biochemistry and Molecular Biology. [Online] Available from:
<https://iubmb.org/publications/standards-for-the-doctoral-degrees/>
[Accessed 23 September 2018]

Jackson, C. and Tinkler, P. (2001) Back to Basics: A consideration of the purposes of the PhD viva. *Assessment and Evaluation in Higher Education*, 26: 4, 355-366.

Johansson, R (2003) Case Study Methodology. *Keynote speech at the international conference: 'Methodologies in Housing Research'* Stockholm, 22-24 September 2003.

Johnston (1997) Examining the examiners: An analysis of examiners' reports on doctoral theses. *Studies in Higher Education*, 22, 333-347.

- Kaestle, C (1993) The awful reputation of educational research. *Educational Researcher*, 22 (1), 23, 26-31.
- Kemp, N., Archer, W., Gilligan, C. and Humfrey, C. (2008) *The UK's Competitive Advantage: The Market for International Research Students*. London: Universities UK, UK Higher Education International Unit.
- Kiley, M (2009) Rethinking the Australian doctoral examination process. *Australian Universities Review*, 51: 2, 32-41.
- Kyvik, S (2014) Assessment procedures of Norwegian PhD theses as viewed by examiners from the USA, the UK and Sweden. *Assessment and Evaluation in Higher Education*, 39: 2, 140-153.
- Lagemann, E (2002) *An elusive science: The troubling history of education research*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.
- Lawrence, S and Sharma, U (2002) Commodification of Education and Academic Labour – using the balanced scorecard in a university setting. *Critical Perspectives on Accounting*, 13 (5-6), 661-677.
- Leathwood, C (2005) Assessment policy and practice in higher education: purposes, standards and equity. *Assessment and Evaluation in Higher Education*, 30 (3), 307-324.
- Lincoln, Y.S., and Guba, E.G. (1985). *Naturalistic Inquiry*. Beverly Hills, California: Sage.
- Lovat, T, Holbrook, A and Bourke, S (2008) Ways of knowing in doctoral examination: How well is the doctoral regime? *Educational Research Review*, 3 (2008) 66-76.
- Lovat, T, Holbrook, A, Bourke, S, Fairbairn, H, Kiley, M, Paltridge, B and Starfield, S (2015) Examining doctoral examination and the question of the Viva. *Higher Education Review*, 47 (3), 5-23.
- Lovitts, B. (2006) in Maki, P. and Borkowski, N. (Eds) *The Assessment of Doctoral Education: Emerging Criteria and New Models for Improving Outcomes*, Sterling, Virginia: Stylus. Chapter 6.
- Lovitts, B. (2007) *Making the Implicit Explicit: creating performance expectations for the dissertation*, Sterling, Virginia: Stylus.
- Maki, P. (2006) Introduction. In Maki, P. and Borkowski, N. (Eds) *The Assessment of Doctoral Education: Emerging Criteria and New Models for Improving Outcomes*. Sterling, Virginia: Stylus. Introduction.
- Miles, M and Huberman, A (1994) *Qualitative Data Analysis: an expanded sourcebook (2nd edition)*. Thousand Oaks, California; London: Sage.
- Morley, L (2004) Interrogating doctoral assessment. *International Journal of Educational Research*, 41 (2), 91-97.

Morley, L, Leonard, D and David, M (2002) Variations in Vivas: quality and equality in British PhD assessments. *Studies in Higher Education*, 27: 3, 263-273.

Mullins, G and Kiley, M (2002) 'It's a PhD, not a Nobel Prize': How experienced examiners assess research theses. *Studies in Higher Education*, 27(4), 369-386.

Murray, R (2009) *How to Survive your Viva: Defending a thesis in an oral examination (2nd edition)*. Maidenhead: Open University Press, McGraw Hill Education.

Murray, R (2011) *How to Write a Thesis (3rd edition)*. Maidenhead: Open University Press.

NCIHE (1997) *Higher Education in the Learning Society*. London: Her Majesty's Stationery Office.

Nightingale, P (1984) Examination of Research Theses. *Higher Education Research and Development*, 3:2, 137-150.

Noble, K. (1994) *Changing Doctoral Degrees: An International Perspective*. Buckingham: Society for Research into Higher Education and Open University Press.

Nyquist, J. (2002) The PhD: A Tapestry of Change for the 21st Century. *Change*, November/December 2002, 13-20.

Nyquist, J and Woodford, B (2000) Seven Propositions from 2000 Conference. Seattle: University of Washington, 2000.

OECD (1996) *The Knowledge-based Economy* [Online]
Available from: <https://www.oecd.org/sti/sci-tech/1913021.pdf>
[Accessed 25 January 2017]

Orpheus (2012) *Standards for PhD education in biomedicine and health sciences in Europe* [Online] Available from:
<http://orpheus-med.org/images/stories/documents/ORPHEUS-AMSE-WFME-standards-for-PhD-education.pdf>
[Accessed 23 September 2018]

Oxford Economics (2017) The economic impact of UK universities 2014-15. Report for Universities UK. [Online] Available from:
<https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2017/the-economic-impact-of-universities.pdf>
[Accessed 23 September 2018]

Park, C (2003) Levelling the playing field: Towards best practice in the doctoral viva. *Higher Education Review*, 36(1), 47-67.

Park, C (2005) New variant PhD: The changing nature of the doctorate in the UK. *Journal of Higher Education Policy and Management*, 27, 2: 189-207.

Park, C (2009), *The Research Student Experience: lessons from PRES*. York: Higher Education Academy.

- Parsloe, P (1993) Supervising Students for Higher Degrees by Research in a Social Work Department. *Journal of Further and Higher Education*, 17, 3: 49-60.
- Pearce, L (2005) *How to examine a thesis*. Maidenhead: SRHE and Open University Press.
- Phillips, E and Pugh, D (2010) *How to get a PhD: a handbook for students and their supervisors*. (5th edition) Buckingham: Open University Press.
- Poole, B (2015) Examining the doctoral viva: perspectives from a sample of UK academics. *London Review of Education*, 13, 3, 92-105.
- Pring, R (2004) *Philosophy of Educational Research* (2nd edition). New York; London: Continuum.
- Pring, R (2014) *Critical Realism: finding a middle way*. Paper presented at the International Conference for Critical Realism, London.
- Pring, R (2015) *Critical Realism: the middle way between idealism and nominalism*. Paper presented at the International Conference for Critical Realism, Indiana.
- QAA (1999) Code of Practice for the assurance of academic quality and standards in higher education. *Section 1: Postgraduate Research Programmes*. Gloucester: QAA
- QAA (2003) *Learning from Subject Review 1993-2001*. Gloucester: Quality Assurance Agency for Higher Education.
- QAA (2007) *Report on the review of research degree programmes: England and Northern Ireland*. QA 116. Gloucester: QAA.
- QAA (2008) UK Quality Code for Higher Education; Part A: Setting and Maintaining Academic Standards. *The framework for higher education qualifications in England, Wales and Northern Ireland* [Online] Available from: <https://www.qaa.ac.uk/docs/qaa/quality-code/qualifications-frameworks.pdf> [Accessed 23 September 2018]
- QAA (2011) UK Quality Code for Higher Education, Part B Assuring and Enhancing Academic Quality. *Chapter B11: Research Degrees*. [Online] Available from: <https://www.qaa.ac.uk/quality-code/the-existing-uk-quality-code/part-b-assuring-and-enhancing-academic-quality> [Accessed 23 September 2018]
- QAA (2012) *Outcomes from Institutional Audit: 2009-11 Postgraduate Research Students*, Third Series. Gloucester: Quality Assurance Agency for Higher Education.
- Radder, H (2010) in Radder, H (Ed) *The Commodification of Academic Research: Science and the Modern University*. Pittsburgh: University of Pittsburgh Press.
- Roberts, G (2002) Set for Success: The Supply of People with Science, Technology, Engineering and Mathematical Skills. HM Treasury. [Online] Available from: http://webarchive.nationalarchives.gov.uk/20091115201121/http://www.hm-treasury.gov.uk/ent_res_roberts.htm [Accessed 23 September 2018].

- Russell, C (Professor, The Earl) (1988), Letter to the Editor, *Times Higher Education Supplement*, 26 February 1988, p.12.
- Rust, C (2002) The impact of assessment on student learning. *Active Learning in Higher Education*, 3 (2) 145-158.
- Schwab, K (2014) (Ed) The Global Competitiveness Report 2014-15 [Online]
Available from: http://www3.weforum.org/docs/WEF_GlobalCompetitivenessReport_2014-15.pdf
[Accessed 07 November 2017]
- Scott, D and Morrison, M (2005) *Key Ideas in Educational Research*. London: Continuum.
- Shuell, T. (1986) Cognitive conceptions of learning, *Review of Educational Research*, 56 (4) 411-36.
- Slavin, R (2002) Evidence-Based Education Policies: Transforming Educational Practice and Research. *2002 Dewitt Wallace-Reader's Digest Distinguished Lecture*, presented at the annual meeting of the American Educational Research Association, New Orleans, 02 April 2002.
- Stake, R (1995) *The Art of Case Study Research*. Thousand Oaks: Sage.
- Stake, R (2006) *Multiple Case Study Analysis*. New York: The Guildford Press.
- THE (2015) *PhD: is the doctoral thesis obsolete?* [Online]
Available from:
<https://www.timeshighereducation.com/features/phd-is-the-doctoral-thesis-obsolete/2020255.article>
[Accessed 23 September 2018]
- Tinkler, P and Jackson, C (2000) Examining the Doctorate: institutional policy and the PhD examination process in Britain. *Studies in Higher Education*, 25 (2) 167-180.
- Tinkler, P and Jackson, C (2002) In the dark? Preparing for the PhD viva. *Quality Assurance in Education*, 10 (2) 86-97.
- Tinkler, P and Jackson, C (2004) *The Doctoral Examination Process*. Maidenhead: Society for Research into Higher Education and Open University Press.
- Trafford, V (2002) Questions in a doctoral viva: views from the inside. Conference paper. UKCGE Research Degree Examining Symposium, Regent's College, London, 29 April 2002.
- Trafford, V (2003) Questions in doctoral vivas: Views from the inside. *Quality Assurance in Education*, 2003, 11:2 114-122.
- Trafford and Leshem (2008) *Stepping stones to achieving your doctorate: focusing on your viva from the start*. Maidenhead: Open University Press, McGraw Hill Education.
- UKRI (2018) Terms and Conditions of Research Council Training Grants [Online]
Available from: <https://www.ukri.org/files/legacy/news/training-grants-january-2018-pdf/>
[Accessed 25 June 2018]

Universities UK (2013) *Patterns and Trends in UK Higher Education* [Online] Available from: <http://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2013/patterns-and-trends-uk-higher-education-2013.pdf>
[Accessed 10 November 2017]

Universities UK (2017) *Patterns and Trends in UK Higher Education: Table 2* [Online] Available from: <http://www.universitiesuk.ac.uk/facts-and-stats/data-and-analysis/Documents/patterns-and-trends-2017.pdf>
[Accessed 10 November 2017]

Vitae (2018) What do doctoral graduates do? [Online]
Available from: <https://www.vitae.ac.uk/doing-research/are-you-thinking-of-doing-a-phd/what-do-doctoral-graduates-do#Data>
[Accessed 09 April 2018]

Wagner, L (1993) The Teaching Quality Debate. *Higher Education Quarterly*, 47(3), 274-85.

Walker, G, Golde, C, Jones, L, Conklin Bueschel, A and Hutchings, P (2008) *The Formation of Scholars: Rethinking Doctoral Education for the Twenty-First Century*. San Francisco: Jossey-Bass.

Wallace, S and Marsh, C (2001) Trial by ordeal or the chummy game? Six case studies in the conduct of the British PhD viva examination. *Higher Education Review*, 34:1, 35-59.

Ward Schofield, J (2000) Increasing the Generalizability of Qualitative Research. In Gomm, R, Hammersley, M, and Foster, P (2000) (Eds). *Case Study Method: Key Issues, Key Texts*. London: Sage.

Watson, D (1995) Quality Assessment and 'Self-regulation': the English Experience, 1992-1994. *Higher Education Quarterly*, 49:4, 326-340.

Wellington, J (2013) Searching for 'doctorateness'. *Studies in Higher Education*, 38:10 1490-1503.

William, D (2004) Standards: what are they, what do they do and where do they live? In Moon, R, Ben-Peretz, M and Brown, S (2004) (Eds). *Routledge International Companion to Education*, London: Routledge. Chapter 20, 351-363.

Wolcott, H (1994) *Transforming Qualitative Data: Description, Analysis and Interpretation*. Thousand Oaks: Sage Publications.

Wood, E and Vella, F (2000) IUBMB updates PhD Standards. *Bioessays*, 22, 771-773.

Yin, R (2009) *Case Study Research Design and Methods (4th edition)*. Los Angeles: Sage Publications.

Appendix 1
Case Summaries
A to J

Analytical framework

This framework emerged from the observations and was used to analyse the notes from each viva.

1 Welcoming introductory remarks to put candidate at ease, some including finding out what the candidate plans to do next (N), others commenting on how much they liked the thesis (code for telling the candidate they would pass as long as the viva was satisfactory?)

2 Explanation about the viva process and what will happen (V) / what outcomes will be (O)

3 Examiners mentioned the importance of checking candidate had written the thesis

4 How and why topic was chosen

5 Leading role in questioning taken by internal (I), external (E), or equally (=), the latter usually being that a different examiner leads different lines of questioning or that they approach the viva collaboratively.

6 Invitation to summarise their most important or significant findings / what they are most proud of in their research

7 Invitation to talk about anything the candidate might have done differently:

B = at the beginning; M = during the viva; E = towards the end

8 Degree of specificity about section of thesis being discussed:

Ch = chapter; P = page; S = section heading

9 Examiners helpful in questioning, enabling reflective approach by candidate and prepared to re-frame questions to aid candidate's understanding of what they are probing

10a Methodology – rationale and execution:

R = why the candidate made their methodological choices

T = theoretical perspective and whether the study is appropriately situated in the field

S = use of statistical methods and design of experiments

RM = discussion of research methods

10b Methodology: length of discussion in proportion to viva overall:

(S) = short; (M) = medium; (L) = long

11 Questions about data and how it was generated and analysed/interpreted

12 Suggestions for changing or augmenting the conclusions

13 Detailed discussion of findings: what candidate thinks they have achieved and what further research is indicated by the study that might be carried out either by the candidate or by others

14 Formative assessment during the viva, relating to: improving the thesis (T), publications emerging from the thesis (P) (including advice about specific content for publication) and general encouragement to publish (E); additional literature to include in the review (Li); how to re-present data for publication (SP); the possibility of using material for lectures (Le)

15 Opportunity for the candidate to raise something the examiners have not asked about

16 Length of viva:

- a) From beginning to when candidate and observer leave for examiners to discuss or until there is a break in proceedings, including time spent by examiners discussing outcome and sharing their decision with the candidate
- b) Total length of viva including examiners' decision-making break

17 Reprising positive comments about the thesis in more detail towards the end of the viva, having completed constructive criticisms

18 Extent to which the viva was overtly testing the candidate's personal and professional skills (PS) as well as the quality and content of the thesis (TQC)

19 Type of corrections:

Mi = minor; Ma = major Re = require candidate to re-write at least one chapter (Ch) or the abstract (Ab) or conclusions (Co)

20 Examiner(s) giving marked-up copy of thesis to candidate to support corrections process:
E = external; I = internal

21 Concluding comments

22 Overall confidence of candidate and ability to defend thesis; attitude of candidate to examiners' suggestions

C=confident

Di=diffident, put off stride

A=accommodating of examiners' suggestions

De=defended thesis effectively when challenged by examiners

Ch=challenged examiners' suggestions

23 Additional comments

Figure 4.3, p.98, shows the positioning of actors in each viva.

Case A Context

Candidate

Broad subject area: Education (AHSS)

Field and qualifications: Assessment practice (MPhil, MSc, PGCE)

Nationality: UK

Mode and stage of study: full-time, mature (31-40)

Submission type: thesis

Examiners

Internal: experienced; candidate's university department

External: professor; experienced examiner and supervisor

Convenor

Not present

Supervisors

Main supervisor: professor; experienced examiner and supervisor

Interviews conducted with: Candidate, external, internal, main supervisor

Length of viva: 1 hour, 35 minutes, including a ten-minute break for finalising conclusions and disclosing recommendations to candidate.

Summary

The viva opened with a brief informal discussion about the candidate's new job and previous experience. The examiners then told the candidate they liked the thesis very much and expected a positive outcome, so they should not be concerned if they 'put them through their paces'. The examiners emphasised it was important the candidate could demonstrate they wrote the thesis.

First the examiners explored with the candidate how the thesis topic was chosen. It emerged from consultation with the supervisor, the candidate having identified that little had been written in the area. Another factor was that the candidate had access to 'live' participants for the study. The examiners asked how a potentially broad topic had been narrowed down to create a manageable PhD research topic, which led to a detailed discussion about the choices the candidate had made in designing the research and focusing on the main research question. During this discussion the external examiner suggested amending the abstract that, as it stood, did not appear to them to do justice to the thesis. This seemed to be an example of examiners wishing to uphold thesis standards.

The leading role in questioning the candidate about methodology was taken by the external examiner, whereas the internal focused more on questions concerning data.

Fairly early on in the viva, the examination appeared to evolve into a conversation with the candidate about the research, constructive criticisms such as 'The data is very rich but you seem to jump to generalising...' made in such a way that suggested peer review.

At the beginning of an extended, detailed discussion about methodology, the external asked the candidate what they were hoping to discover when they began the study. The candidate did not have a hypothesis as such in mind. They explained the rationale behind a mixed methods study. The external probed the candidate's implementation of the study, suggesting

alternative approaches to the research and questioning whether they had adopted a positivist stance and if an ethnographic perspective would have been more appropriate.

Retrospectively, this was designed to ensure the candidate could defend how they had conducted the research and that the methodology and methods were suited to answering the research question. The candidate was able to give a strong defence of their approach to the research, importantly giving reasons for making various decisions during the field work.

One of the strengths displayed by the candidate was a willingness to carefully consider the external's suggestions even if not accepting them, for example, discussing if they should have structured the thesis differently. Similarly, when the internal examiner made some suggestions about the data sets used in the study, the candidate answered thoughtfully and made a good case for what they had done. Similarly to some of the other AHSS vivas, research methods and methodology were a recurring theme.

Towards the end, the examiners asked the candidate about the impact of the study. The candidate's response led to further discussion about how bias had been avoided and any subject-specific outcomes accommodated. The examiners suggested that the study offered various insights into the process under scrutiny and that the candidate should write a journal article on three related elements emerging from the thesis. The candidate agreed this was a possibility, again accommodating the examiners' suggestions. Later, the external returned to the theme of publication, asking the candidate if they had further contributions to make. The candidate had several ideas for publications from the thesis and other data they could not use in the PhD. Both examiners expressed a wish that the candidate would publish from the thesis.

The examiners were enthusiastic in their praise for the thesis quality when disclosing their recommendation for minor corrections to the candidate. They mentioned, for example, that the thesis was: thorough; well written; clear; and made good use of existing research. They added that the quality of presentation was excellent, there had only been a handful of errors and that the substance was also excellent, with a richness of evidence. The candidate had extracted many interesting and original insights. They thought the conclusions could have been longer; however, they were succinct yet gave important insights. The examiners had very much enjoyed reading the thesis because of the ideas and how well written it was. They suggested that the best theses were those with unfinished business.

Finally, the examiners promised to give their marked-up copies of the thesis to the candidate for corrections, congratulated them and confirmed a re-written abstract should be sent to the internal examiner. They would send their joint report to the University's Graduate Studies Committee when corrections had been completed. This appeared to be a strong thesis, well defended by the candidate.

Case B Context

Candidate

Broad subject area: Education (AHSS)

Field and qualifications: International study of teachers in a culturally regulated context (prior qualifications unknown)

Nationality: International

Mode and stage of study: full-time, non-mature

Submission type: thesis

Examiners

Internal: experienced examiner and supervisor,
candidate's university department

External: experienced examiner and supervisor

Convenor

Not present

Supervisors

Main supervisor: experienced examiner and supervisor
from candidate's department

Interviews conducted with: internal examiner, main supervisor

Length of viva: 1 hour, 30 minutes, including a ten minute break for finalising conclusions and disclosing recommendations to candidate.

Summary

The viva opened with warm introductions, the examiners thanking the candidate for a 'fascinating' thesis. The internal examiner asked the external if they would like to begin. The external declined, so the internal asked the candidate what they were most proud of and what went best with the study. The candidate responded that the findings had related well to the literature and this had been striking, given that at first it had been challenging to find relevant studies. The internal next asked if in retrospect the candidate would have done anything differently. One possibility had been to add focus groups to the research methods and to share with them some of the initial findings, but this might have created problems, for example, participants challenging the research outcomes. In response to a question from the external, the candidate suggested that sharing research outcomes with participants would have radically changed the data. Participants may not have wished to admit the evidence and it could have compromised their anonymity. The candidate described the difficulty of communicating the research, even to those involved, because of perceptions of its lack of value. When asked by the external if they wished to undertake further analysis, the candidate agreed they would have liked to explore one of the chapters with participants.

The examiners complimented the candidate on their critique of the literature and contribution to knowledge, although suggested they could have been more authoritative in dismissing some of the literature. The internal questioned the candidate on the extent that two of the authors of particular relevance to the study represented scholarly thinking in the field. This led to a discussion of those authors' work, the candidate asserting they were not lone voices in the political context, followed by a detailed discussion of the literature and how it had affected the candidate's theoretical perspective. The candidate admitted that if writing a paper from the thesis they would focus more on participant identity in changing contexts. The candidate agreed with the examiners' suggestion that they could have been more dismissive of some of

the literature, explaining that the scarcity of publications concerning this topic in the geographical region was linked to a lack of focus on the individual and the difficulties of understanding individuality in a rapidly changing culture. The external commented that the way participant identity had come to the fore in the study was emerging much more in the discussion than it had in the thesis. The candidate agreed that the thesis would have been different if they had been considering the topic from their current perspective, using a socio-cultural approach.

The internal asked the candidate about the extent to which their personal beliefs had had to be set aside when conducting the research. The candidate asserted they had been able to assume the role of objective researcher. They had deliberately chosen a theoretical perspective that would enable them to explore individual opinions without imposing a personal view. The candidate's 'localness' and being from the same culture had helped to convince participants to consent and engage with the study, yet they were still in some senses an outsider. Although difficult, it had been possible to set aside their own beliefs to attain objectivity, helped by their supervisor and the participants' wish to keep some distance. More observations and focus groups would have provided more insight into how participants interacted and to gauge their views as a collective. The candidate would have liked to explore why some of the participants' views had not been translated into practice, but noted that what constituted acceptable innovation varied between institutions. It emerged that the state-sponsored candidate had had to obtain permission to conduct research in public institutions, including explaining why they were not using a questionnaire. The external wondered if the results showed parallels with other groups in similar professions: did others experience tensions between their preferred practice and external expectations? The external asked about an element of the research that was brought into the thesis at the end: the candidate's understanding of the institutional culture and context affecting individual practitioners. This had left the external wondering about the influence of different institutions on individual practice. The candidate had wished to focus on individuals in the study, although had draw from different institutions to achieve diversity in the participant sample.

Combining observations with interviews had led the candidate to particular choices of terminology so as to differentiate between implicit and explicit practice and to identify patterns and cases that fell outside them but were used to corroborate findings. The external agreed that the focus of individuals and their practice had made the study more powerful and that the candidate had made a good case for how the data were analysed. The external asked if the candidate had reviewed their findings to check all participants had been quoted. The candidate had not, but explained that some of the respondents had 'stood out' for their contribution.

The examiner-candidate dialogue was striking for its similarity to a discussion among experts in the field. The candidate was respectful of the examiners' arguments and suggestions, yet defended their approach and the choices made. They accommodated some of the examiners' suggestions, even though they were not disposed to change terminology, and agreed they would do some things differently knowing what they did at the end of the study. Much of the questioning focused on methodology and data analysis in the cultural context of the study. While the examiners were encouraging and positive throughout, they asked the candidate probing questions. The viva resulted in a recommendation for minor corrections. At the end, the examiners congratulated the candidate, praising in particular the participant 'vignettes', which deepened and enriched the work. The final chapter and viva had answered their questions. They hoped the work would be disseminated.

Case C

Context

Candidate

Broad subject area: Biosciences (STEM)

Field and qualifications: Primatology; background in social anthropology (Masters)

Nationality: EU

Mode and stage of study: full-time for 4 years, part-time for 3 months, mature (31-40)

Submission type: thesis

Examiners

Internal: candidate's university department

External: professor based in internationally recognized
UK independent research institute

Convenor

Present

Non-cognate STEM
field

Supervisors

Director of studies/main supervisor and second supervisor: both experienced and from candidate's department

Interviews conducted with: candidate; external and internal examiners; both supervisors

Length of viva: 3 hours, 20 minutes, including a five minute break after 1 hour, 20 minutes, 15 minutes' finalising conclusions, report-writing and disclosing recommendations to candidate.

Summary

At the start, the convenor introduced everyone, explained their role in the process, and informed the candidate they could ask for a break if needed, and would be asked to leave the room at the end of the discussion. The external congratulated the candidate on behalf of both examiners for producing a 'nice' thesis, which was impressive in terms of the quality of work and the findings. They would begin with general questions, then go through the thesis in chronological order.

The internal asked why the candidate had chosen the research questions and to study the troupe of animals. The candidate explained they were interested in how the dynamics of groups work and how individuals fit in; the sociology of groups was important in this field. They chose this troupe, which was already known to be cohesive, because of existing research taking place that would allow them to build on other findings. A measure was needed that would identify subtle differences. The external enquired about the most interesting and important findings from the study. The candidate seemed prepared for this question and answered it without hesitation, providing a full answer that it was the comparison between associated groups of the species and their behaviour, focusing on cohesion and aggression and how tensions were resolved. The external agreed this was a striking point.

Having been asked by the external what they would do differently if starting again, the candidate replied that they would have used a different unit of analysis. They focused on individuals but would have liked to study group behaviour in more detail. The internal enquired if studying more groups would have been helpful. The candidate agreed, adding it would have been helpful if the groups had been a similar size. It emerged that at one stage

both the groups studied had been the same size but the candidate had had to return to the UK to renew their visa and during that period nine individuals had disappeared.

The examiners were explicit in questioning the candidate about the thesis, mentioning page numbers and specific text. For example, the external asked for clarification on whether a period of observation had been continuous or made up of several different periods. Both examiners thought the thesis introduction was good and could be turned into lectures. Page references were also used to initiate discussion about the research design and statistical analysis, to suggest more direct comparison with the work of one author whose research was closely aligned with the study, and to advise more detail was needed throughout. On one occasion the candidate asked the examiners to re-frame a question, which they did successfully.

The external was interested in the time period of observations and how rules for conducting observations had been interpreted, showing great interest in the conduct of the research, enquiring about the behaviour of individual animals in the context of the length of time of observations. The candidate and external discussed the tension between using longer observation times (used by another researcher) and not being able to sample so many animals. Having been asked whether they would double the observation time if conducting more research, the candidate defended the time period used in the study, which they thought satisfactory. The internal asked a detailed question about gender in the context of group behaviour and if the physical approach to observations was the best way to capture the optimum number of subjects observed. The examiners agreed the candidate's approach was rigorous, although questioned the large number of variables involved in the analysis. The candidate defended their choice but was open to the examiners' suggestions. The examiners linked their advice to a possible approach to publications from the thesis.

The examiners asked numerous questions about data generation, analysis and interpretation, which included discussion of subject behaviour and food choice in relation to the relatively low statistical power of the data. The candidate had interpreted the data correctly but could have been more cautious, for example, recognising the relatively small sample sizes and other factors to be taken into account. The examiners suggested the candidate evaluated the statistical power of the test and made it clear the results will be treated with caution where the power was low.

A short break was followed by intensive discussion of data collection and how comparisons across different groups were made, returning to the question of statistical significance. The external suggested the small samples were 'an intrinsic limitation' of the results and that this could be overcome by larger samples. They were particularly critical of one chapter, suggesting it lacked clear hypotheses, despite good research questions. If the candidate intended to publish from the thesis, it would be useful to have some predictions. Both examiners were highly complimentary about the final results chapter, citing coherence and strong methodological approaches. They provided further advice about publications, including mention of putting the candidate in touch with the editor of a new journal.

An extended discussion of the thesis findings and conclusions became an animated dialogue among peers, showing parity of esteem, with the candidate and examiners suggesting further research. The examiners summarised the parts of the thesis they would like the candidate to strengthen. The candidate, who displayed confidence and openness to the examiners' suggestions despite probing questions, passed with minor corrections. After asking if the candidate had anything to add, the convenor explained the next steps in the process.

Case D Context

Candidate

Broad subject area: Media, culture and language (AHSS)

Field and qualifications: International comparative study of media policy and regulation (Masters)

Nationality: International

Mode and stage of study: full-time, mature (41-50)

Submission type: thesis

Examiners

Both examiners were external to the university, due to a lack of expertise in the field internally; both were examiners and supervisors

Convenor

Present
Non-cognate AHSS experienced field

Supervisors

Main supervisor (director of studies) was experienced supervisor; attended the viva; co-supervisor, inexperienced, also attended

Interviews conducted with: candidate, one external, main supervisor and convenor

Length of viva: 1 hour, 20 minutes, plus 45 minutes for finalising conclusions, decision-making and disclosing recommendations to candidate.

Summary

The convenor made introductions, explaining their role was to ensure 'fair play' and adherence to university regulations and emphasising it was important the candidate felt confident and able to provide the best possible defence. If the candidate needed a break, they should ask. The viva was intended to be as helpful as possible and not confrontational. If the candidate had no questions (they did not), the viva could begin. The examiners had difficult questions for the candidate, but hoped that they would nevertheless enjoy the exam. They took turns in pursuing different lines of questioning, appearing to have agreed this in advance and demonstrating a co-operative approach.

The examiners asked why the candidate had embarked on the research: what was their motivation and background? When the candidate had summarised their professional background and the context for the study, the examiners asked them to explain their thesis, seeking clarification of why the three particular countries were chosen for comparison. One suggested the candidate's thesis was difficult to sustain, referencing a thesis page and data used by the candidate to discuss the extent of public ownership in the media sector. In response, the candidate referenced a chapter where they had presented a rationale for choice of the countries. Having complimented them on the 'push and pull' approach to addressing research questions, both examiners challenged the candidate's assertion that the three countries were comparable, suggesting they were at different stages of development in media policy. The candidate had attempted to show all three had mixed systems. When the examiners queried a key element of the candidate's methodology, suggesting that they might have been expected to 'look at what others had done' in framing this research and that they were seeking a detailed description of how the correlation was undertaken, the candidate replied they were constrained by the thesis word limit and did not have time or space to explain this. The examiners also queried why no 'elite' interviews were conducted. The

candidate agreed these would have added to the richness of data and the cultural impact of the thesis but there had not been enough time within their 3-year completion period, or sufficient space in the thesis. They had decided to prioritise analysis of other data and did not think the absence of elite interviews had compromised research outcomes. The examiners remained convinced that elite interview data and reference to other literature would have supported the candidate's claims for convergence and enabled the thesis to be more embedded in the field. Formative feedback included suggestions for shortening some chapters to exclude some of the technical material to allow the inclusion of more analysis. The candidate was asked what changes they would make to the thesis, on reflection. They referred to introductory chapters that, as the examiners had suggested, did not contribute to their argument. Discussion returned to the inter-country comparison and different media contexts. The candidate referenced the conclusions to support their arguments, growing in confidence as the viva progressed. They strongly articulated the practical, policy and political implications of their findings and emphasised the industrial, social and cultural benefits.

The examiners had 'specific questions' about the conclusions, presaging corrections they would recommend. They wondered why there was new information in the conclusions, suggesting this was surprising. It was intriguing for them to find a mixture of normative and analytic approaches and that the normative. The candidate explained their rationale for including the policy recommendations that had emerged from their analysis which were there for 'ease of reference'. They agreed that what the examiners perceived as new information was 'a bit of a bizarre point'. The examiners confirmed this information belonged somewhere else in the thesis, offering suggestions for what to emphasise. One was 'totally bemused' about some of the material in the conclusions.

Rather than moving chronologically through the thesis, the examiners referenced chapters and pages to develop points they wished to explore. The thesis appeared to require some restructuring and revision, with the findings constituting a contribution to knowledge. Both examiners were courteous, e.g. giving the candidate time to reflect before answering questions, while challenging their methodology and methods. They asked the candidate about their career ambitions and if they were considering publishing the thesis as a book, minus the introductory chapters. At the end, the examiners asked if the candidate wished to ask a question or mention anything not already covered. The candidate asked what they thought of the thesis, at which point the convenor explained the examiners would inform the candidate of their recommendations and to them as an outsider it seemed to have been a fair discussion. Both examiners wanted to be sure the candidate had had enough opportunity to tell them about the research. Both thought the conclusions could be clearer and more focused and emphasised this was not unusual and not a bad thing for the candidate to reflect having completed the research. The candidate stressed that everything in the conclusions had been mentioned earlier. The examiners told the candidate they had worked 'incredibly hard' and done 'a fantastic amount of work'. It was always difficult to know what to leave in and what to take out of the thesis.

The examiners recommended the award of a PhD subject to minor corrections, to be completed within three months. These comprised: re-writing the abstract, a clear statement of the contribution to knowledge, making the thesis more analytical and re-writing conclusions to do the study justice, as well as correcting spelling errors, tidying up references and reproducing tables and graphs in colour. The examiners had enjoyed meeting the candidate and reading the thesis, offering many congratulations on the positive result. Their joint report would be emailed to the candidate and the Director of Studies would oversee the corrections, which they hoped would not take long.

Case E

Context

Candidate

Broad subject area: Sociology (AHSS)

Field and qualifications: Women's studies (Masters; academic background in FE)

Nationality: UK

Mode and stage of study: part-time, mature (51 or over)

Submission type: thesis

Examiners

Convenor

Internal: experienced; candidate's university department; Not present

External: professor; experienced examiner and supervisor

Supervisors

Main supervisor: professor, now working in another university; experienced supervisor and examiner; attended the viva

Co-supervisor: experience unknown, not interviewed

Interviews conducted with: candidate and internal

Length of viva: 1 hour, 35 minutes, including 15 minutes for finalising conclusions and disclosing recommendations to candidate.

Summary

The examiners thanked the candidate for attending and explained the Department required them to audio record the viva in case of any problem and that the candidate was welcome to a copy. Both examiners were very impressed with the thesis and enjoyed reading it. They explained that they had both general and specific questions on the thesis and that at the end, they would ask the candidate to step outside while they considered the outcome. The external began by asking the candidate how they got involved in the topic. The candidate responded that they had been interested in it for a long time, mentioning an important author in the field and how their professional background involving different roles in the field had played a part in the choice of topic.

The internal examiner asked the candidate to outline the original contribution to knowledge in the thesis. The candidate referred to three strands of the research, each of which they believe contributes to knowledge in the field, including methodological originality. This led to a detailed general discussion of the study, beginning with a question from the internal about areas the candidate had not been able to include. The candidate mentioned three themes that had been omitted, either because of lack of space in the thesis or because they would make another PhD study. This response led to a discussion of the sector the candidate had researched, including the notion of 'service' in employment in this context. The external queried an apparent lack of humour in the participants, which the candidate assured them had not been completely the case and gave an example of a humorous situation that had occurred. However, the context of the observations had not lent itself to a great deal of humour. Early in the discussion, the candidate referred to potential family influence on the behaviour of the participants, referring to their first names in order to be specific.

The external complimented the candidate on how well they had written about the topic, from the perspective of an 'insider/outsider' relationship, which had varied depending on whether

the participant was in one of the two roles that had been observed for the study. The external identified that the candidate's relationship with one group appeared more closed than with the other group. The candidate explained why this was the case and in which situations they had been an outsider or insider and some of the benefits of feeling like an outsider in the different situations observed. It was important not to feel too integrated when taking a critical approach. This had been challenging in one example in particular when an interviewee spoke particularly openly to the candidate as if they were part of the same group. This became a peer discussion, the external complimenting the candidate on writing 'nicely' about the tensions inherent in the situation and suggesting the candidate had been hard on themselves. Considerable further discussion took place concerning the methods and methodology, the internal querying whether the candidate had given feedback to the two groups. This had not yet happened but the candidate wished to do this because the participants had been so generous with their time, but the problem was that they had not been very aware of what was behind the research and what the candidate's main interest was. To overcome this, feedback might be given on how the differences between types of learning can be over-emphasised.

Changing direction, the external was intrigued by the candidate's theoretical perspective, which they perceived as a 'bizarre mix' of three theories. The candidate explained how their theoretical perspective had developed, citing several well known theoreticians in the field. The external thought the candidate had provided a good defence of an eclectic choice of theoretical perspective, asking the candidate a further question about practical implementation and congratulating them on the quality of argument. Discussion of the findings included a peer discussion on the situations of the study contributors and their capacity to be critical of the sector they were working in. They appeared to have few individual choices; some were aware of and angry about the power relationships involved in their work. The external suggested the participants' unrealistic career aspirations were exacerbated by recent government policy, their inability to acknowledge reality resulting from their bleak situations. The candidate agreed that participants were making a virtue of necessity. There followed a three-way discussion about the role of class in the debate and its influence on the situation of the participants.

Both examiners gave formative feedback concerning the work of other authors, even though they did not necessarily suggest referencing them in the thesis. They also gave advice about publishing from the thesis and undertaking further research. The examiners had finished their questions and asked if the candidate wished to ask them any questions. The candidate requested further advice about publishing.

After a 15-minute break, both examiners told the candidate they had been very impressed with both the thesis and the defence of it. They would be recommending the candidate be awarded a PhD subject to minor corrections. They advised them to ask someone to proof-read the thesis before finalising it. They would provide a list of corrections for the candidate. The procedure was that the internal examiner would submit forms to the examinations office with their recommendations. Their report would be considered by the University's examinations committee, who would then write to the candidate confirming they had 28 days in which to make changes. A revised copy of the thesis should be sent to the internal examiner. At the end, the external congratulated the candidate on 'one of the best defended theses I have ever examined', suggesting they really knew what they had done and why. This was interesting from the point of view of the observer: throughout much of the viva the candidate had appeared rather diffident, yet responded strongly to challengers from the examiners and grew in confidence.

Case F Context

Candidate

Broad subject area: Film studies (AHSS)

Field and qualifications: Documentary film-making (Masters)

Nationality: UK

Mode and stage of study: part-time, film-maker; part-time lecturer; mature (41-50)

Submission type: Thesis by published work; 1-year registration; two documentary films accompanied by a supporting statement

Examiners

Internal: experienced; candidate's university department;

External: professor and international expert in the field with extensive experience as an examiner and supervisor

Convenor

Present

Non-cognate AHSS field

Supervisors

Main supervisor was experienced supervisor and examiner

Attended the viva

Co-supervisor: experience unknown, not interviewed

Interviews conducted with: candidate, external, internal and main supervisor

Length of viva: 2 hours, 20 minutes, including 1 hour for finalising conclusions, report writing and disclosing recommendations to candidate.

Summary

The examiners complimented the candidate on the work at beginning of viva and gave the impression of being impressed with the candidate's films, mentioning the words 'unique' and 'captivating'. Their admiration of the practical element of the submission was also evident in their interview responses. This was not a 'traditional' thesis but a practitioner submission that constituted two documentaries and a supporting statement. The external examiner confirmed in their interview that the candidate's films clearly demonstrated a contribution to knowledge; however, the examiners required them to re-write elements of the supporting statement to bring it into alignment with comparable PhD submissions in the field.

The examiners did not enquire about the candidate's future plans until the end of the viva, but in responding to their questions at the beginning, the candidate mentioned dual interests: film-making and film education, and passing on knowledge to students. This was a mature candidate whose career as a film-maker would continue in parallel with an academic role. The academic element of the candidate's future plans may have influenced the corrections the examiners requested.

The candidate's preparedness for the viva was influenced by guidelines sent by the University to make candidates aware of how the viva would be conducted. They had experienced a mock viva with both supervisors that helped them to prepare for the real thing. However, the candidate pointed out that the supervisors' feedback on the supporting statement was different from the questions asked by the examiners, so the opportunity to practise was useful for knowledge of format but not content. In their interview, the candidate expressed surprise about some elements of the viva but had felt able to defend everything written in the supporting statement. The candidate's perspective was supported by the observation: the

convenor explained the viva process to the candidate at the beginning of the examination and the candidate appeared to defend the thesis confidently, while accepting examiners' suggestions. In summary, this was a well-prepared candidate who confirmed having enjoyed the viva experience.

The case demonstrated multiple examples of feedback to the candidate, both to improve the supporting statement and to support them in their academic career, including advice on publications. The positioning of the research in the wider field was discussed at length and included: (i) a book the candidate had written and referenced in the literature review but which s/he had not been permitted to include in the PhD submission since it was not an academic publication; (ii) advising the candidate to include a wider range of publications in the literature review (and to give full acknowledgement to other authors); and (iii) the theoretical approach taken by the candidate and the coherence of the decision to compare two works seen by the examiners as rather disparate. These three items of feedback suggest a greater emphasis on preparing the candidate for an academic career rather than professional practice; however, they were also focused on the thesis quality and its implications.

Discussion of methodology and theoretical perspective included exploration of the candidate's decision to undertake research in two unfamiliar environments, as an 'outsider' and the question of whether the theoretical debate about this approach should have been mentioned in the thesis. The examiners continued to question the candidate about the theoretical approach throughout the viva, with the discussion suggesting the complexities of integrating two works of visual art with a supporting statement to produce a PhD submission by published work. The examiners suggested that the clarity with which the candidate responded to some of their questions should be reflected in the thesis. The candidate's professional skills as a film maker were fully tested in the viva but the examiners' approach suggested a primary interest in the academic coherence and quality of the submission and that it should reflect the candidate's research abilities.

At several points in the viva, especially when the candidate was responding to examiners' questions about how the data were generated and when the examiners were suggesting that one element of the supporting statement was worthy of independent publication, the discussion appeared to the observer to have progressed to a conversation among subject specialists. The examiners demonstrated a deep interest in the topic, complimenting the candidate, on 'incredible' use of archive material. This case illustrated the continuum of judgement inherent in the thesis-viva relationship, where the viva ended in a discussion among peers.

The viva lasted longer than in some of the other AHSS cases, indicating the complexities involved in examining a PhD by published work and deciding on recommendations. At the end, the external examiner confirmed that the examiners would be recommending the award of PhD with minor corrections. The convenor confirmed the candidate should re-submit no longer than three months after the viva and that the internal examiner would sign off the corrections. The candidate was required to re-write both the introduction to the supporting statement and the literature review, and to integrate the overall structure so as to replace the 'fragmentary nature of the written statement' and to make more of the archive section.

Case G

Context

Candidate

Broad subject area: Biosciences (STEM)

Field and qualifications: Animal behaviour; predator-prey reactions (BSc)

Nationality: International

Mode and stage of study: full-time, non-mature (teaching experience as GTA)

Submission type: thesis (including chapter already published as journal article)

Examiners

Internal: very experienced examiner from another college in the university; not topic expert

Non-cognate STEM field

External: Less experienced examiner; expert in candidate's field

Supervisor
Experienced supervisor
from candidate's dept.

Convenor

Present; from another college in the university

Interviews conducted with: candidate; external and internal

Length of viva: 2 hours, 45 minutes, including finalising conclusions and disclosing recommendations to candidate.

Summary

The examiners introduced themselves when the candidate entered and explained the purpose of the convenor. The candidate had prepared a presentation of their findings but the examiners declined this. In the pre-viva discussion, the examiners had discussed how they wished to proceed and as there was complete agreement about the thesis, they took a joint, co-operative approach to questioning. As the opening question of the viva, the internal asked what the candidate thought was the most interesting result, what made them proud of the work? The candidate referred to the first of two experiments and explained why they had chosen the method and why the experiment had been particularly successful in its design, leading to a new finding about decision-making in this species.

Both examiners had liked the introduction, but the external suggested the candidate had not explained why it was important to study the topic. The introduction needed more information about the species being studied, including basic behaviour, physiology and capability. The internal advised the candidate to remove a statement implying a causal relationship that could not be proved and to avoid such statements if they could not be justified by scientific evidence. The internal then advised the candidate that: the examiners would go through the thesis chapter by chapter; this was a conversation amongst adults; and that there was plenty in thesis that made it good enough to pass, so they should relax, even though some changes would need to be made. The candidate thanked the examiners, who then asked about the candidate's background (in biology). The PhD involving predator-prey reactions was a natural progression from the BSc. The candidate explained their plan was to undertake genetic research in a lab and that they had secured a post-doctoral position in another country, working on the same species. The internal informed the candidate they would take the viva very seriously; it really mattered that the thesis was 'right', since they were entering a scientific career.

Throughout, the examiners made explicit reference to chapters and pages in the thesis. The viva contained much formative feedback, especially on the use of statistical methods. Both

examiners were exercised about the lack of information concerning authorship of the chapter that had already been published in a journal. It was not clear how many individuals had contributed to the work, or the extent of the candidate's participation. The candidate asked for advice about how to present the chapter. The convenor intervened to confirm that the candidate was required to be completely clear about their contribution and what others had done. There was some discussion about how the chapter had come to be published and how the journal had originally returned the paper for further work after review, before accepting it. The discussion highlighted the challenges inherent in including prior publications in the thesis, e.g. a need to remove any ambiguities and ensure there was sufficient information for the reader to make sense of it in a different context. The examiners offered further advice about how the candidate should present the publication as a thesis chapter, confirming that having already published was in the candidate's favour.

Highly detailed discussion took place of the candidate's research methods and decisions taken during complicated experiments, e.g. why had the candidate used a particular measurement rather than others. The candidate explained these choices, which the examiners were happy with, but would have liked the explanation to be in the thesis. A justification was needed for settling on the measurement and an explanation about the software used. Much of the examination focused on the candidate's statistical analysis. In particular, during an intensely detailed discussion of analysis of the findings, the examiners advised against trying to show something was statistically significant when this was not the case, providing much formative feedback to the candidate on what statistical tests would be valid for their data and how the results should be interpreted and presented. They intended that all this formative assessment would enable the candidate to conduct more accurate tests for statistical significance that would produce evidence to support statements made in the thesis. During this discussion the candidate asked the examiners many questions about conducting and presenting statistical tests and their results, going into much detail, e.g. even querying how many decimal points to use. It was striking how receptive the candidate was to the examiners' advice, which was given openly, constructively and obviously to increase the candidate's knowledge and understanding. The examiners explained the experiments were impressive in their design, yet the statistical analysis meant the candidate had not done the results justice. Knowing the candidate's career plans influenced the examiners' particular attention to the statistics. They explained that as the candidate was planning to become an academic scientist, they needed specific professional skills, including understanding how to conduct statistical tests and to be able to analyse and present the results accurately. Discussion of the findings was dominated by further formative feedback concerning statistical tests and presentation of the results. The examiners suggested that statistical weaknesses meant the candidate had not used enough of the data, had restricted what was being tested and had not made the most of the richness and variation in the data.

Exploration of the findings from the second experiment continued to include formative feedback, e.g. adjusting the experiment design for future use, although showed elements of peer discussion, including possibilities for further research. The examiners also advised the candidate to increase references to the wider field, especially similar experiments. At the end, the examiners congratulated the candidate, confirming they were impressed with the research. The candidate had passed with major corrections, to be completed within nine months, as the examiners wished changes to be made to each chapter. The external handed their annotated copy of the thesis to the candidate to help with corrections. The candidate thanked the examiners profusely for all their constructive feedback. The examiners would provide references to other publications the candidate might find useful.

Case H

Context

Candidate

Broad subject area: Biosciences (STEM)

Field and qualifications: Immunology (2 x BSc)

Nationality: EU

Mode and stage of study: full-time, mature (26-30); work experience in a laboratory)

Submission type: thesis (including work to be published as journal article)

Examiners

Internal: experienced examiner from candidate's university department

External: professor based in UK teaching hospital research laboratory

Convenor

Not present

Supervisor

Experienced supervisor from candidate's dept.

Interviews conducted with: candidate; external and internal

Length of viva: 3 hours, 15 minutes, including 5-minute break, finalising conclusions and disclosing recommendations to candidate.

Summary

The viva opened with positive comments from the external examiner who had much enjoyed reading the thesis, which they described as 'interesting', 'well-written' and 'easy to read'. When asked at the beginning to summarise the most important findings arising from the research, the candidate's response included an intention to develop the research in a post-doctoral position in North America; therefore the examiners were aware from the start that the candidate was aiming for a research career. This knowledge may have affected the examiners' feedback. The candidate's preparedness was striking in that they were able, when interviewed, to summarize precisely the characteristics and abilities they thought examiners would be expecting them to display. The viva held no surprises for them. The internal examiner recognised this: '*I was very impressed during the viva...[the candidate] was very good actually...unfazed and...many of the questions we asked [they] had obviously thought about before*'. When the candidate left the room during a 5-minute break, the examiners agreed they were defending the thesis well.

From the outset the examiners' attention to detail in the thesis was evident from their allusions to various figures and pages and the specificity of questioning. The nature of the candidate's topic necessitated in depth discussion of a range of types of cells, their behaviour in different situations, and the impact on the cells of different substances present in the body, e.g. some proteins and their contribution to a person's immunity, or lack of it, to specific diseases, in this case rheumatoid arthritis. Discussion of the results of the candidate's experiments and interpretation of the results were therefore the focus of the viva. The examiners stated they would 'go through' the thesis: this meant a chronological and comprehensive discussion of each chapter. The approach suggested both examiners had read the thesis with great care and in minute detail to understand the research and its outcomes and to analyse the candidate's findings. During the viva observation the internal examiner displayed knowledge of the candidate's topic, i.e. the behaviour of cells and immunology in rheumatoid arthritis, while the external examiner was experienced in the field of cell biology with a particular interest in immunology, statistical techniques and their interpretation. This

balanced combination of expertise together with the candidate's confident defence of the thesis brought an intensity and equality to the discussions in the viva, much of which could have been an exchange of views among peers rather than an examination.

This case contained examples of explicit feedback, with the examiners focusing on asking the candidate to explain how experiments were conducted and to interpret the findings in detail. The depth of the candidate's knowledge of the topic was such that in one example, they had found papers relating to a particular immunology therapy that the internal (the topic expert) had not heard of; in another, the candidate's hypothesis on the behaviour of some cells led the internal to ask if s/he was planning to publish this (although the candidate wished to do further work first). However, the examiners drew the candidate's attention to the importance of being clear about ownership. The candidate's response to one of the internal examiner's questions suggested that part of the thesis introduction was due to be published as a jointly authored journal paper, currently awaiting reviewers' comments. In discussion, it emerged that experimental data from a clinical trial that had taken place in their department in parallel with the candidate's own experiments, had been included in the thesis. The external asked why the candidate had not excluded those results, the reason being that it would have been difficult to separate them from the other work in the study and that viewed together the research raised important questions. The candidate explained that the molecular work for the PhD study was undertaken with a post-doctoral researcher working in the candidate's university department. The candidate knew how the experiments were done but the post-doc, who had more experience of molecular biology, had carried out most of the work because it was quicker and they were close to submitting the paper. As a result of this exchange the external suggested that the candidate might add an explanation about how this element of the work had been done; it was important to be clear about one's own work, especially now collaborative research was so prevalent.

In discussing generation, analysis and interpretation of data, the external examiner questioned the candidate closely on the statistical analysis and interpretation of the data, with both examiners providing constructive suggestions for interpretation and the external confirming the candidate had built 'a solid case'. The external also asked the candidate about problems encountered when trying to grow a large number of cells (of a type difficult to culture) that had been lost due to technical issues, suggesting this had been bad luck as it was a large sample. This line of questioning suggested an interest in how the candidate had dealt with complications in the research; the confident response provided by the candidate appeared to show this problem had been successfully overcome. Throughout, the candidate appeared to possess sophisticated problem-solving skills when responding to technical questions and when the external asked what would s/he take forward from the research the candidate had already thought of further research to test the hypothesis. This viva was one of the longest observed, resulting in a recommendation for minor corrections, mainly concerning the presentation of statistical data in the thesis.

Case I Context

Candidate

Broad subject area: Biosciences (STEM)
Field and qualifications: Primatology (BSc, MRes)
Nationality: UK
Mode and stage of study: full-time, mature (31-40)
Submission type: thesis

Examiners

Internal: Experienced examiner from
candidate's department

Convenor

Present
Non-cognate STEM

External: Professor, experienced
Examiner

Supervisor

Experienced supervisor
from candidate's dept.

Interviews conducted with: candidate; both externals and supervisor

Length of viva: 3 hours, 10 minutes, including 20 minutes for decision-making, report writing and disclosing the result.

Summary

The examiners had enjoyed reading the thesis, which was good and of a standard they expected for doctoral work. It was long but easy to read and clearly written. The candidate made their points clearly and had a good writing style. The external asked the candidate to take five minutes to 'talk us through': why this topic and what was the most significant finding.

This species was of interest because they were successful and adaptable – the candidate was attracted by the flexibility of the species, both in their forms of communication and in their diet. These were the main areas of interest, but the candidate was particularly fascinated in the more complex elements of communication: language, syntax, and symbolism, for example, and different ways of vocal and tactile communication in the wild. The candidate demonstrated they were aware of a large body of work concerning flexibility in communication, but they wished to explore why this was so fundamental to more complex elements of communication in the wild. The candidate also explained in detail their most exciting finding concerning communication within families.

The first external took the opportunity to ask further questions during the explanation and the viva appeared to take the form of peer discussion at this early stage, the second external expressing surprise at some of the findings. The candidate mentioned the work of other researchers and the first external referred to their annotated copy of the thesis, which they would give the candidate at the end. The examiner mentioned that 'we' had noted minor comments, some of which might be useful for publications, suggesting both examiners had agreed about feedback and changes to the thesis. The examiners then informed the candidate they would go through the thesis in some detail but did not intend to raise every point.

Both examiners showed great interest in the candidate's research, especially the methodology, methods and how the results had been interpreted and analysed. The candidate had identified interesting behaviour, which was not part of the study but planned to undertake further

research on this in the future. The examiners and the candidate discussed in great detail how data had been collected in the field. This included the candidate describing the long periods of time spent preparing to collect data and how the pilot study had helped to determine the final configuration, which had consisted of both observed and unobserved time. Both examiners questioned the candidate about the precise locations of the species when data was collected and how they had relied on hearing at times when they could not observe the animals. The candidate had found it easy to identify individuals from hearing and had used acoustic analysis, including recordings of different calls made to substantiate their own hearing.

The first examiner offered detailed formative feedback on presentation of the findings, which the candidate appreciated. The candidate had chosen not to classify a particular type of behaviour as intentional due to lack of evidence in the study, whereas the examiner suggested others would have judged it intentional. The examiner referred to papers they had published with another researcher that might help the candidate to approach the findings slightly differently. The second examiner questioned the candidate about the amount of data collected compared with the length of time spent, suggesting that they would have expected more data to have been produced. The candidate explained the challenges they had faced in ensuring integrity of the data, including taking time to recognise particular individuals. The first examiner had also experienced some of the same challenges and was sympathetic.

The first examiner asked the candidate to explain one of the graph types used. When the candidate had done this, the examiner was satisfied but asked them to provide a simpler explanation that would link the graph to other graphs. The format was neat but it needed to be explained. Continuing the discussion of presentation of data, the first examiner complimented the candidate on a 'fabulous job' of avoiding 'pseudo-replication' but thought that at one point this had occurred. The candidate thought not because they had conducted further statistical tests for this data. The examiner explained why they had thought the candidate had pseudo-replicated the data and was gracious in apologising for misunderstanding. Both examiners complimented the candidate on the 'very nice' chapter being discussed and suggested it showed a model that could be replicated by other researchers.

Other chapters were also discussed in detail, in particular the extent of reliability of the data and contextual factors the candidate did not control for in their statistics and that might have affected reliability. During discussion of the research, the candidate displayed thorough knowledge of the topic and the field and an excellent grasp of their statistical analysis, including how they established which variables were significant or not. Much of the viva took the form of a peer discussion. Both examiners wished to explore in great detail with the candidate the behaviour of the individuals studied and what deductions had been made in different circumstances. They appeared to thoroughly test the candidate's research capabilities and some professional skills, in particular interpretation of statistics and communication of the results.

At the end of the viva, the convenor asked the candidate if they wished to add any information about the thesis, or if they had questions or comments for the examiners. The candidate declined. After the 20 minute break, the examiners congratulated the candidate: 'We were really impressed'; 'We were particularly impressed with the thoroughness of the stats'; and 'A very impressive piece of work, very worthy of a PhD'. The candidate passed with minor corrections, to be completed within 3 months.

Case J - Context

Candidate

Broad subject area: Dentistry (STEM)

Field and qualifications: Microbiology (Bachelor of Dental Science; Masters in orthodontics)

Nationality: International

Mode and stage of study: full-time, mature (practising dentist) (35-40)

Submission type: thesis

Examiners

Internal: experienced examiner from cognate university department

External: professor from UK university

Convenor

Not present

Supervisor

Experienced supervisor from candidate's dept.

Interviews conducted with: candidate; external and internal; supervisor

Length of viva: 3 hours, 15 minutes, including 30 minutes for decision-making, and disclosing the result.

Summary

This viva had a somewhat unusual beginning: on entering the room the candidate handed a revised graph to the external examiner, to replace one of the figures in the thesis. The external thanked the candidate, who now had an opportunity to tell the examiners about the background to their study and why they had chosen this topic for their PhD. After graduating as a dentist in their home country, the candidate had studied orthodontics overseas for one year as a postgraduate. Returning home, they practised as a dentist before coming to the UK to undertake the PhD, while practising part-time. Their practice led to their interest in implants and periodontal disease and the wish to develop an effective treatment.

It had taken four years to design the protocol for the study. The candidate had had the idea of looking at different bacteria and treatments, without knowing about microbial analysis or in vitro research. They had sent their proposal to the university, and the professor who later became their main supervisor helped to revise it. The research took three and a half years, most of which had been spent in the laboratory. The candidate was self-funding but their department had paid for the last five months of their third year. They began with two supervisors but when some of the microbiologists in the department had been made redundant around 18 months into their programme, they had to learn how to manage the microbial data on their own. When problems arose with their experiments, there had been no one to ask so they had to approach other students and another microbiology department for help. During this period the main supervisor had 'disappeared', just as the candidate was starting to write up. They managed to get advice regarding data analysis from another academic. Another challenge they faced was in finding sufficient patients who matched the inclusion criteria for the study, although eventually they tested 21 patients. They had wanted 25 but could not increase the sample and analyse data in the time available.

The external advised the candidate that the examiners would go through the thesis and ask about some of the things they had not understood. The internal established that the candidate was aware of who both examiners were. The internal first asked how much of the work the candidate had done. The candidate had done everything (sequencing, taking samples) except the clinical protocol for ethics purposes, which the main supervisor had done. The candidate

explained that a hygienist had taken samples, but that they had also been there to measure them and do x-rays. The candidate undertook all the analysis, taking samples to the lab and abandoning non-parametric tests after being advised to normalise the data. The external asked the candidate about help they had received to write the thesis. While their main supervisor had provided general advice on the thesis structure, the candidate had realised from reading other theses that they were structured differently from theirs (e.g. including an introduction and general discussion), so had added a preface. The external asked the candidate if they viewed the thesis as three distinct studies. The candidate did not, but they thought it might be confusing to have a general discussion at the end. The external advised the candidate the examiners would go through the three findings chapters in detail. Anything they did not mention, the candidate should assume was fine.

The external commented on the lack of general information in the thesis concerning the field of implants and its history. The candidate thought this might be irrelevant and that the examiners might be more interested in why periodontal disease occurs. The external explained that while the examiners were aware of the history, when the thesis was published, other readers may not know the general context, which needed to be added, to include, for example, an explanation about why titanium remained the metal of choice for implants. What was the candidate's understanding of this? The candidate provided an in-depth explanation of previous research involving titanium implants and what had been discovered, alluding to other metals that were not as compatible. The examiners explained the need for the candidate's research to be set in the context of the wider field, and that it was important they understood the extent of the candidate's knowledge, for example, why the surface properties of different metals affected disintegration.

The examiners proceeded to discuss the candidate's research and results in great depth, exhibiting great interest in the work, the findings and their application, while demonstrating there thesis would require significant amendment. This was a peer discussion that enabled the candidate to justify their approach and explain the implications of their findings, such as the wider risks associated with periodontal disease. It emerged that the candidate had done much more work than was evident from the thesis, which the internal suggested needed to be referenced. The internal probed the candidate's knowledge of the literature, suggesting additional references of relevance to the study. There was a lengthy discussion about statistical analysis and data presentation and although the examiners were content with the statistical tests and the candidate's interpretation of their significance, they asked for clarification of some figures. The external was impressed the candidate had done all the sequencing and asked if they had had any technical support. A lab technician had taught the candidate what to do and also helped when something went wrong. The internal remarked on the 'beautiful photos' showing stages of the disease. Both examiners displayed a highly supportive approach, realising the candidate had faced challenges and had worked partly in isolation. They were rigorous in their questioning, however, and despite their enthusiasm for the research, discussed with the candidate a long list of revisions to improve the thesis. Nevertheless, the examiners agreed the candidate had presented much useful experimental data. The candidate, while accepting the shortcomings of the thesis, defended their research and findings with confidence, displaying knowledge and understanding of the field not apparent from the thesis. The viva concluded with congratulations and praise for the candidate's major achievement in producing high quality data. They should be proud of this and their hard work. Nevertheless, the thesis would require major amendments, to be completed within six months if possible, which the examiners had already spelled out but which would be in their written report and would make the thesis easier to read. The examiners recognised the challenges the candidate had faced and did not want them to be disheartened. Most PhDs had corrections and the candidate had coped well. Well done.

Appendix 2

Information sheet and consent form for participants

Provided to participants on University of Oxford, Department of Education, headed paper
November 2012

INFORMATION SHEET FOR PARTICIPANTS

Title: The final assessment of the doctorate: a study of how examiners assess PhD candidates

This form asks for your consent to participate in a D Phil research study about how the PhD is assessed. If you would like further details about the study to help you decide whether you would like to participate, please contact me at the email address above.

The purpose of the research is entirely for my D Phil thesis. I am a self-funding, part-time, mature student with practical experience of doctoral education and long-standing involvement in policy development for research degree programmes, especially the doctorate. My interest in the assessment of the PhD arises from this professional experience. The results of the study will be published in my doctoral thesis, which will be publicly available on the Oxford University research archive website. Results may also be disseminated in academic publications and presented at conferences.

By agreeing to participate, you are agreeing:

- i) either to allow me to observe a viva in which you are taking part, as a candidate, examiner, or as the supervisor of the candidate and to be interviewed by me after the viva has taken place; or
- ii) to be interviewed by me as an examiner or supervisor of multiple PhD candidates, unrelated to a particular viva.

The consent options are set out in the Participant Consent form attached. Interviews will be recorded unless you decline this by deleting the text in italics.

Benefits and risks: Your participation in this study will contribute to knowledge about the final assessment of PhD candidates (thesis, or equivalent, and viva). Any benefits to you will depend on your interest in the topic. Confidentiality of responses will be maintained; all the information collected as part of this study will be used anonymously. Care will be taken not to enable identification by a third party of any university or individual involved in the study. I will be storing data as password-protected files on my personal lap-top which is not accessible by anyone else. The data generated by the study will be coded and stored anonymously, although I will keep a hard copy of all the information in case it is necessary for any reason in the future to be able to track responses and individuals (for example, should you wish to withdraw at a later date – see below). The hard copy information will be stored in a locked cupboard in my home to which only I have access and will be destroyed at the end of the study. Your interview recording will also be erased at the end of the study; were you to decide to withdraw from the project at any point, your recording would be erased immediately. All participants will receive a copy of the conclusions of the study for information and I hope that the outcomes will be of interest and use to individual examiners and supervisors, as well as to universities.

Withdrawal from the project: If at any time after consenting to taking part in this study you change your mind, for whatever reason, you can withdraw immediately without having to give me a reason. Should you wish to do so, please email me as soon as possible and I guarantee to

remove any data you have provided from the data archive. There will be no penalty for withdrawal.

Use of the data: If, once my DPhil thesis is publicly available, I am approached by other researchers who wish to re-use my data, I will only release anonymised information to genuine researchers.

Complaints: Should you have a complaint of any kind about this study, please contact: Professor Ingrid Lunt, Director of Doctoral Programmes, Department of Education, Oxford University, at the above address or by email: Ingrid.Lunt@education.ox.ac.uk

<p>This project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee.</p>

PARTICIPANT CONSENT FORM

1) Participants

Please tick the relevant box(es):

i) Interviews relating to viva observations

Candidate:

I am a PhD candidate and agree to Gill Clarke observing my viva examination

I also agree to being interviewed after the viva *and to the interview being recorded**

Examiners:

I am a PhD examiner and agree to Gill Clarke observing a PhD viva that I am examining

I also agree to being interviewed after the viva *and to the interview being recorded**

Supervisors:

I am the PhD candidate's supervisor attending the viva and agree to Gill Clarke being present as an observer

I also agree to being interviewed after the viva *and to the interview being recorded**

ii) Interviews unrelated to viva observations

Examiners and Supervisors:

I am a PhD examiner and/or supervisor and agree to being interviewed by Gill Clarke about the final PhD examination , *and to the interview being recorded**

** Please delete text in italics if you do not wish the interview to be recorded*

I agree voluntarily to participate in this study. I have read and understood the terms and conditions and the nature of the research described in the Information Sheet. I have had the opportunity to ask questions about the study and have received satisfactory answers to any questions.

Name _____

Title _____

University _____

Signature _____

Date _____

2) Researcher

I agree to abide by the terms and conditions set out in the Information Sheet for participants

Name _____ Gillian Clarke _____

Title _____ D Phil student _____

University of _____ Oxford _____

Signature _____ *Gillian Clarke* _____

Date 1st November 2012 _____

Appendix 3

Exploratory letter to institutions

Provided to institutions on University of Oxford, Department of Education, headed paper

June 2011

Dear

PhD / DPhil in Education: Assessment of the PhD

I am a part-time, first year research student (DPhil) in the Department of Education at Oxford University, supervised by Professor Ingrid Lunt. I am writing to ask whether your University might be prepared to take part in my doctoral research project, outlined below.

The working title of the project is: *The final assessment of the doctorate: a study of how examiners assess PhD candidates*. The study aims to explore the qualities and characteristics that examiners of doctoral candidates are looking for when they assess doctoral candidates in the final examination and how they go about making their judgements. I am interested in the reference points examiners use, taking account of their previous examining experience, subject conventions, any criteria and regulations provided by the university, and external guidance. The project will include both components of the PhD final examination, namely, evaluation of the thesis or equivalent (for example, portfolio or artefact), and assessment of the candidate in the *viva voce*, and will explore the weighting given to the two components.

I recognise that this is a sensitive and complex area in which to undertake research and intend to design and carry out the project in a way that acknowledges and allows for this. The current research design, which is still developing, relies on my having access to PhD examiners' reports and on asking a sample of examiners and supervisors to complete a questionnaire and/or, if they are agreeable, to be involved in follow-up interviews. Ideally I would also like to be able to interview doctoral graduates after their viva, but the design and methodology of the study are flexible enough to omit this step if it proves too difficult to arrange.

Subject comparisons are part of the study and at present the planned disciplines to be surveyed are biosciences, chemistry or physics, economics, civil engineering, dentistry or medicine, law, modern languages and philosophy. These subjects have been chosen because they represent a range of disciplines and are frequently found in UK universities, although not all the institutions involved in the study will offer all of them.

A limited amount of empirical research is available about the final assessment of the PhD in the UK in different subjects, and particularly about how judgements are made in practice about the two components of the final examination, including the criteria and reference points used by examiners. The study would allow detailed qualitative information to be sought from examiners, supervisors and candidates which might be of interest to others in these roles.

The purpose of contacting you at this stage is to ask whether in principle you would give permission for me to have access to examiners' reports and staff, as outlined above, on the understanding that **all information collected as part of this study will be used anonymously and care will be taken to ensure that the universities and individuals involved are not identifiable**. You would also receive a copy of the conclusions of the study. Should you wish to have further information about how the study will be conducted before deciding whether to participate, I would be happy to provide it. To test the feasibility of the research design and potential willingness to participate, only a few universities are being contacted now, to ask whether they would be prepared to pilot the study. Eventually I hope that up to 20 different

institutions from across the UK will be involved and to include data gathered from the pilot universities in the final analysis.

Your willingness to be involved in this study would be much appreciated. I can be contacted by email or phone as indicated above and look forward to hearing from you.

Yours sincerely,

A handwritten signature in black ink that reads "Gill Clarke". The signature is written in a cursive, slightly slanted style.

Gill Clarke

Appendix 4
Interview schedules:
supervisors and examiners; candidates;
and independent chairs/convenors

Interview schedule for supervisors and examiners

1 Criteria used to make judgements in the final PhD assessment and how they relate to outcomes

Questions for supervisors

1.1 Assuming the candidate has met all the progress hurdles in previous years, how do you know when to advise him/her that s/he is ready to submit for the PhD assessment? On what evidence do you base your judgement?

1.2 Before the final assessment, do you take a view about whether a PhD candidate is:

- Exceptional?
- Good enough to pass without major corrections?
- Borderline or likely to be asked to make major corrections?

If you do have these or similar categories in mind, how often do examiners' reports and the assessment result differ from yours? Are you ever surprised by the outcome of the final exam? How are these points relevant to the final assessment of the candidate just examined?

1.3 In your view, what attributes/characteristics/abilities/skills are examiners seeking in PhD candidates? From your experience are these likely to be similar in different disciplines? How is your answer borne out by the candidate just examined?

Questions for examiners

1.4 When reading the candidate's work (thesis or equivalent) as part of a final PhD assessment, do you take a view about whether the candidate is:

- Exceptional?
- Good enough to pass without major corrections?
- Borderline or likely to be asked to make major corrections?

What evidence do you use to make such a judgement? How are these points relevant to the final assessment of the candidate just examined?

1.5 If you do have these or similar categories in mind, to what extent were your pre-viva impressions borne out at the end of the final assessment you have just completed?

1.6 How often do you find that your co-examiners' views, pre- and post-viva, are similar to your own? Was this the case in relation to the candidate just examined? As examiners, what steps do you take before the viva to establish that you are both/all using similar criteria?

1.7 Are you ever surprised by the final outcome of the PhD assessment? Were you surprised by the outcome of the assessment just completed?

1.8 As an examiner, what attributes/characteristics/abilities/skills are you seeking in PhD candidates? For example, what questions did you have in mind when considering the recent candidate's work (thesis or equivalent) and during the viva? Did you benchmark the person with other candidates you've examined? To what extent, if at all, did you have in mind any

external criteria (including guidance at subject level)? When answering this question you may want to think about the assessment just completed. From your experience are these likely to be similar in different disciplines?

1.9 How and when during the PhD assessment just completed did you come to a final judgement about the candidate? Was this timing similar to other candidates you've examined?

1.10 What do you think are the respective roles of the internal and external examiners, if they are different?

2 Outcomes of the final PhD assessment

2.1 Please comment on the comparability or diversity of final PhD assessment outcomes (thesis and performance in the viva) among doctoral candidates in your subject or broad field, and if possible comparing the candidate just examined with others.

2.2 The current final PhD assessment is at threshold level - the candidate either passes or fails. Do you think the PhD should be graded? Why / why not?

2.3 Whether or not you think the PhD should be graded, should exceptional performance in the final assessment be rewarded in the outcome? You might think this would apply in the viva just completed? If your answer is 'yes', how could/should this be done, for example, a formal award of merit/distinction such as a starred PhD that would indicate excellence? If so, should this be recorded on the degree certificate?

2.4 In your experience, approximately how many / what percentage of candidates fail the final examination? 0%, <3%, 4-10%, >10%

2.5 In your experience approximately how many PhD / what percentage of candidates are (a) awarded a lower degree or (b) asked to complete major corrections (sometimes called 'referral')? 0%, <3%, 4-10%, >10%

2.6 Do you think fewer or more candidates should fail the final examination and / or be awarded a lower degree? Please give reasons.

2.7 If you think you have enough cross-disciplinary experience, can you please say whether you think the current final PhD assessment leads to comparable levels of achievement by PhD graduates across different subjects?

2.8 How does the final assessment of a PhD candidate enable all those involved to be confident that the right decision has been made about the outcome? In answering this question, please draw directly on the final assessment you have just completed, as well as others.

2.9 How much, if at all, are growing numbers of doctoral candidates affecting the process and outcomes of the final PhD assessment?

3 Weighting of the components (thesis and viva) in the final outcome

3.1 What weight do PhD examiners give to the two components (thesis or equivalent and viva) of the final PhD assessment? What was the weighting in the assessment you have just completed? In your opinion, does this vary across subjects / fields?

3.2 In your experience, does the judgement about one of the components ever change the outcome of the other part of the assessment, and/or of the final result? In answering this question, please think about the assessment you have just completed.

3.3 What do you think is the purpose of the viva? Again, please draw on the examination just completed.

Interview schedule for candidates

1 Criteria used to make judgements in the final PhD assessment and how they relate to outcomes

1.1 Having met the progress hurdles (e.g. transfer and confirmation of status), during your degree programme, how did you know when you were ready to submit your thesis for the final PhD assessment?

1.2 Before the final assessment of your thesis and before the viva, what did you think you would have to do to be awarded a PhD? For example, on what basis did you think examiners would be making their judgements, about your thesis and about you as an individual researcher? What formal guidance, if any, are you aware of that suggests what examiners should take into account? And what attributes/characteristics/abilities/skills do you think examiners are looking for in PhD candidates?

1.3 Assuming that doctoral examiners' judgements are already based on firm foundations using well established and accepted criteria, can you think of any way(s) in which the final PhD assessment could be different in order to increase the likely integrity / correctness of PhD assessment outcomes?

2 Outcomes of the final PhD assessment

2.1 You are answering these questions following your viva: to what extent are your pre-viva impressions concerning what you would need to do to be awarded a PhD and what examiners are looking for, borne out now you have experienced the final assessment?

2.2 Are you surprised by the result of your final PhD exam? Why?

2.3 Do you think the PhD should be graded? Why / why not?

2.4 Whether or not you think the PhD should be graded, should exceptional performance in the final assessment be rewarded in the outcome? If your answer is 'yes', how would you prefer this to be done, for example: congratulatory letter or prize; or formal award of merit/distinction? Should any of these be recorded on your degree certificate?

- 2.5 How does the final assessment of a PhD candidate enable all those involved to be confident that the right decision has been made about the outcome?
- 2.6 How much, if at all, do you think that growing numbers of doctoral candidates are affecting the process and outcomes of the final PhD assessment?

3 Weighting of the components (thesis and viva) in the final outcome

- 3.1 The thesis and viva together are what you are judged on in the final PhD assessment in the UK. What weight do you think PhD examiners give to each? Do you think that either is more important? Are the two parts equally necessary for evaluating you and your research?
- 3.2 Did you notice any variation between the examiners in the weighting they seemed to give to the two components?
- 3.3 What do you think is the purpose of the viva?
- 3.4 Was there any indication from your experience that your viva had changed the examiners' opinion and potentially the final outcome of your PhD assessment overall?
-

Interview schedule for independent chairs/convenors

1 Criteria used to make judgements in the final PhD assessment and how they relate to outcomes

- 1.1 How do you think examiners decide whether a candidate has or has not achieved what is necessary to be awarded a PhD (thesis and viva)? What questions do they have in mind? What attributes /characteristics /abilities/skills do you think examiners are looking for in the candidate? From your experience are these qualities likely to be similar in different disciplines? Do examiners benchmark the person with others they have examined? How do they make their judgement(s)? When answering this question you may want to think about the assessment just completed.
- 1.2 How and when during the PhD assessment process related to the viva just completed do you think the examiners came to a final judgement about the candidate? Was this timing similar to other candidates whose vivas you've chaired?

2 Outcomes of the final PhD assessment

- 2.1 Please comment on the comparability or diversity of final PhD assessment outcomes (thesis and performance in the viva) among doctoral candidates whose vivas you have chaired or convened, if possible comparing the candidate just examined with others.
- 2.2 The current final PhD assessment is at threshold level - the candidate either passes or fails. Do you think the PhD should be graded? Why / why not?
- 2.3 Whether or not you think the PhD should be graded, should exceptional performance in the final assessment be rewarded in the outcome? You might think this would apply in the viva just completed? If your answer is 'yes', how could/should this be done, for example, a formal award of merit/distinction such as a starred PhD that would indicate excellence? If so, should this be recorded on the degree certificate?

2.4 In your experience as an independent chair or convenor, approximately how many / what percentage of PhD candidates fail the final examination outright? 0%, <3%, 4-10%, >10%

2.5 In your experience as an independent chair or convenor, approximately how many / what percentage of PhD candidates are (a) awarded a lower degree or (b) asked to complete major corrections (sometimes called 'referral')? 0%, <3%, 4-10%, >10%

2.6 Do you think fewer or more candidates should fail the final examination and / or be awarded a lower degree? Please give reasons.

2.7 If you think you have enough cross-disciplinary experience, can you please say whether you think the current final PhD assessment leads to comparable levels of achievement by PhD graduates across different subjects?

2.8 How does the final assessment of a PhD candidate enable all those involved to be confident that the right decision has been made about the outcome? In answering this question, please draw directly on the final assessment you have just completed, as well as others.

2.9 How much, if at all, are growing numbers of doctoral candidates affecting the process and outcomes of the final PhD assessment?

3 Weighting of the components (thesis and viva) in the final outcome

3.1 What weight do you think PhD examiners give to the two components (thesis or equivalent and viva) of the final PhD assessment? What was the weighting in the assessment you have just completed? In your opinion, does this vary across subjects / fields?

3.2 In your experience, does the judgement about one of the components ever change the outcome of the other part of the assessment, and/or of the final result? In answering this question, please think about the assessment you have just completed.

3.3 What do you think is the purpose of the viva? Again, please draw on the examination just completed.

4 Role of the Independent Chair or Convenor

4.1 Did you receive formal guidance about your role based on the University's expectations and requirements for independent chairs / convenors?

4.2 If you did receive guidance, how has that influenced the way in which you approach your role? How does the University assure itself that you and other independent chairs /convenors fulfil the role in a similar way?

4.3 What, in your view, is the purpose of the role of independent chair / convenor? Does your own view fit with your University's expectations of the purpose? In responding to this question, please focus on the viva just completed.

4.4 How, if at all, do you think the independent chair /convenor influences the outcomes of the viva examination?

4.5 What do you think are the respective roles of the internal and external examiner in the final PhD assessment, and particularly in the viva? Are there any differences in how they approach the viva?

4.6 How, if at all, would you augment / change the role of the independent chair / convenor as a result of your own experiences?

Appendix 5
Coding framework

The PhD examination process		
Long-code	Short-code	Interview questions
Role-thesis Thesis primary evidence for judgement Has-candidate-done-enough-for-PhD Does-thesis-meet-examiners-core criteria Initial impressions Take-view-re-three-categories (exceptional, minor corrections, borderline) Benefit-of-doubt Benchmark-with-others Peer-judgement Contribution-thesis-final-judgement Thesis-as-indicator-final-outcome	Role-T T-primary-evid-judge Has-cand-done-enough Does-T-meet-core-crit Initial-imp Take-view-3-cat Benefit-doubt Bench-others Peer-judge Contrib-T-fin-judge T-indicator-fin-outcome	
Purpose-viva: candidate Rite-of-passage Possible-plagiarism Check-candidate-understanding Examiners-approach-to-viva: candidate Purpose-viva: examiner/observations To-be-sure-candidate-has-done-work Confirm-judgements-from-thesis Check-candidate-understanding Check-candidate-background What-went-wrong (e.g. lab problem) Ask-about-rationale To-moderate—borderline-thesis Provide-candidate-with-feedback Add-missing-information Avoid-perpetuating-mistakes Improve-statistical-knowledge Influence-of-candidates-next-role Examiners-give-candidate-notes Advice-re-publication: candidate Face-to-face-important Ask-candidate-to-add-anything-important Viva-change-outcome For-better For-worse	Purp-V:C Rite-of-pass 1 Poss-plagiar 2 Check-cand-underst 3 Ex-app-to-V:C Purp-V:E/obs To-be-sure-cand-work 1 Conf-judge-from-T 2 Check-cand-underst 3 Check-cand-background 4 What-went-wrong 5 Ask-rationale 6 Moderate-border-T 7 Prov-cand-feed 8 Add-miss-info 8.1 Avoid-perpet-mistakes 8.2 Imp-stats-know 9 Influ-cand-next-role 10 E-cand-notes 11 Adv-re-public:C 12 F-to-f-imp 13 Ask-cand-add-any-imp 14 V-change-outc 15 For-better 15.1 For-worse 15.2	C-2.1 C-3.3 E-3.3

Viva-does-not-change-outcome Align-examiners-judgements	V-not-change-outc Align-E-judge	16	
Relative-importance-thesis-viva Candidate Examiner-weight: candidate Examiner Set-viva-bar-low-if-excellent-thesis %-relative-importance Poor-thesis-good-viva Good-thesis-poor-viva Most-weight-thesis	Rel-imp-T-V Cand E-weight-C Ex V-bar-low-if-exc-T %-rel-imp Other-prop Poor-T-Good-V Good-T-Poor-V Most-weight-T	1 2 3 4 5 6 7 8 9	C-3.2 E-1.9

Candidate attributes sought by examiners		
Long-code	Short-code	Interview questions
Examiners-criteria: candidate	Ex-crit:C	C-1.2
Originality-contribution	Origin-contrib 1	E-1.4
PhD-is-a-process: candidate	PhD-a-process: C	
Examiners criteria: other factors	Ex-crit: other fact	
Institution-supervisor-reputation	Inst-sup-reput 1	
Peer-reviewed-publication(s)	Peer-rev-public 2	
Candidate-background	Cand-backgr 3	
Progress-hurdles-passed	Progr-hurdle-pass 4	
Examiners-rely-on-supervisor-judgement-thesis-ready	Ex-rely-super-judge-T 5	C-2.5 E-1.8
Supervisor-reputation: candidate	Super-reput-C	
Examiners criteria: initial impressions	Ex-crit: initial imp	
Broad-categories	Broad-cat 1	
Quality-of-thesis	Qual-T 2	
Is-thesis-well-written	Is-T-well-writt 3	
Includes-aims-objectives	Incl-aims-obs 3.1	
Typographical-errors	Typos 4	
How-is-data-presented	How-data-present 5	
Are-statistics-ok	Are-stats-OK 6	
Is-work-original	Is-work-orig 7	
Volume-of-work	Vol-of-work 8	
Importance-of-introduction-in-setting scene	Imp-intro-set-scene 9	
Benchmark-with-others	Bench-w-others 10	
Peer-judgement	Peer-judge 10.1	
Thesis-as-indicator of final outcome	T-ind-final-outcome 11	
Examiners'-overall-criteria-thesis-and-viva	Ex-overall-crit-T+V	
How-much-is-work-their-own-or-supervisor's	How-much-own-work 1	
Candidate-self-critical	Cand-self-critic 2	
Can-ID-weaknesses-in-work	Can-ID-weak-work 2.1	
What-would-they-do-differently	What-do-differ 2.2	
Able-think-critically-at-high-level	Able-think-critic-high-lev 3	
Example-transferability-PhD-skills	Eg-transferabil-PhD-skills 4	
Analyse-data	Analyse-data 4.1	
Make-decisions	Make-dec 4.2	
Critique	Critique 4.3	

Present-work	Present-work	4.4
Able-to-discuss&process-info	Able-disc&proc- info	4.5
Competent-capable-scientist:	Compt-capble-sci	5
Can-they-run-their-own-lab	Can-run-own-lab	5.1
Intellectual-rigour-and-integrity	Intellect-rigr-integr	5.2
If-not-competent-independent- scientist:	If-not-comp-ind-sci:	6
Have-they-completed- enough-training	Compl-enough-trg	6.1
Have-they-enough-skills-for- award	Enough-skill-award	6.2
How-has-supervisor-affected-thesis	How-super-affect-T	7
Has-candidate-read-literature	Has-cand-read-lit	8
Use-of-controls-in-scientific-work	Use-of-controls-sci	9
Does-candidate-understand-own-work	Cand-underst-own-wk	10
Does-candidate-acknowledge-others- contribution	Cand-ack-others-contrib	11
Honesty-and-integrity	Honesty-and-integrity	12
Correct-statistical-analysis	Correct-stats-analys	13
Data-interpretation	Data-interpretation	14
What-would-be-next-stage-of-research	What-next-stage-R	15
Identify-key-findings	ID-key-find	16
Volume-of-data	Volume-data	17
Thesis-content-different-from-paper	T-cont-diff-paper	18
Examiners-supervisors-interest-in- topic: candidate	Ex-sup-inter-topic: C	19
Where-is-threshold	Where-threshold	
Rigour-of-transfer-and-impact-on-progress- to- final-exam	Rig-transf-impact-progr-final- exam	

Examination outcomes: comparability, diversity and grading		
Long-code	Short-code	Interview questions
Range of achievement How-many-candidates-fail How many-candidates-referred-or-lower-degree Should-more-candidates-fail-or-get-lower-degree Comparison-with-standards-in-other-countries Question-of-judgement More-masters	Range-achieve How-many-cand-fail How-many-cand-ref-lower More-cand-fail-lower Comp-stand-other-countr 1 Q-of-judge 2 More-M 3	E-2.4 E-2.5 E-2.6
Potential-outcomes Institutional-guidelines-options Resubmission Minor-corrections Major-corrections-referral Timing-options Candidates-should-not-fail-at-viva-stage	Potent-outcome Inst-guide-opt 1 Resubmiss 2 Minor 3 Major-refer 4 Timing-opt 5 Cand-shd-not-fail-at-viva 6	
Grading: Grading-yes: candidate Grading-no: candidate Grading-yes: examiner Differentiate-above-threshold Grading-no: examiner Judged-on-publications (sci) PhD-process-and-set-of-skills Grade-inflation	Grading: Grad-Y:C 1 Grad-N:C 2 Grad-Y:E 3 Differ-above-thres Grad-N:E 4 Judge-pubs-sci 4.1 PhD-proc-skills 4.2 Grade-infl 4.3	C-2.4
Distinction Distinction-yes: candidate Distinction-no: candidate Distinction-yes: examiner Nominate-for-institutional-prize Distinction-no: examiner	Distinction Dist-Y:C 1 Dist-N:C 2 Dist-Y:E 3 Nom-inst-prize 3.1 Dist-N:E 4	C-2.4

Appendix 6

Extracts from university research degree regulations

1 Guidance on the conduct of the examination, in particular the viva

Some universities include statements in their regulations that show the overall judgement of the candidate rests on both the thesis quality and the performance of the candidate in the viva, provide detailed guidance on its conduct, or summarise its purposes, as illustrated by these extracts from several universities' regulations:

Lancaster University

3. PURPOSE OF THE ORAL/VIVA VOCE EXAMINATION

The oral examination, or viva voce ("by live voice" examination) is a key part of the examination process and allows the examiners to explore with the student areas of interest, controversy, weakness, obscurity, etc., in order that they may then make an appropriate recommendation to the university. All research students must have a viva, regardless of the examiners' views on the quality of the thesis. In some cases the viva may do no more than confirm a favourable opinion already formed from the thesis; in other cases it will provide the candidate with the opportunity to compensate for weaknesses in the thesis or dispel examiners' reservations; in a minority of cases it may fail to do this and instead confirm the examiners' view that a recommendation other than approval should be made. It is, therefore, an important occasion for which students should be as well prepared as possible. Discussion may focus very specifically on sections of the thesis, research methodology, lab work, etc, or may expand to look more generally at the field and the place of the thesis's contribution within it. In summary, the viva voce examination allows:

- the student to defend his/her research and thesis, expanding or clarifying sections, providing more detailed background information, etc.;
- the examiners to obtain further clarification on the work undertaken and the content of the thesis;
- the examiners the opportunity to assure themselves that the thesis and its underlying research are the work of the student himself/herself.

University of Leicester

Purpose and Format of the Viva Voce Examination

Literally, "viva voce" means by or with the living voice - i.e., by word of mouth as opposed to writing. So the viva examination is where you will give a verbal defence of your thesis.

Put simply, you should think of it as a verbal counterpart to your written thesis. Your thesis demonstrates your skill at presenting your research in writing. In the viva examination, you will demonstrate your ability to participate in academic discussion with research colleagues.

Purpose of the Exam

The purpose of the viva examination is to:

- demonstrate that the thesis is your own work
- confirm that you understand what you have written and can defend it verbally
- investigate your awareness of where your original work sits in relation to the wider research field
- establish whether the thesis is of sufficiently high standard to merit the award of the degree

for which it is submitted

- allow you to clarify and develop the written thesis in response to the examiners' questions

University of London

The examiners, after reading the thesis, shall examine the candidate orally and at their discretion by written papers or practical examination or by both methods on the subject of the thesis and, if they see fit, on subjects relevant thereto.

University of Nottingham

Note: the University's Quality Manual: Viva Voce Examinations relies on the QAA doctoral qualification descriptor, other than for guidance on the purpose of the viva as set out below, so by default encompasses the criterion of originality or a contribution to knowledge.

6 Purpose of Viva

The viva will normally include **questions** designed to ascertain that the thesis embodies the candidate's own research. It will test the candidate's general comprehension of the field of study within which the subject of the thesis falls. It will test the candidate's acquaintance with the general literature of the subject, knowledge of the relation of the work to the wider field of which it is a part, and the respects in which the work advances, modifies or otherwise affects this wider field of scholarship.

University of Reading

Assessment during the viva

During the viva, the examiners need to reassure themselves that the candidate understands his or her work, how it fits into the wider literature, what is original about it, what has been found, and what potential impact this may have. They also need to check that candidates are aware of any limitations of their approach and how these might be addressed in future work.

It is not essential for examiners to ask questions about every detailed aspect of the thesis. The usual approach is to ask some general questions in the initial phase of the discussion, and then to work through the thesis, often chapter by chapter. Areas of potential concern will need to be probed in more depth. Having worked through any major issues, the examiners will often adopt a 'sampling' approach in relation to more minor queries (bearing in mind how long the viva has lasted so far).

University of the West of England

K16.5 Examination arrangements

The examination process K16.5.1R

The examination shall have two stages:

the candidate's submission of the thesis / collection of published works and the examiners' independent preliminary assessment of it; and

the defence of the thesis/collection of published works by the candidate by viva voce

or approved alternative examination.

K16.5.2R A candidate shall normally be examined by viva voce examination on the research covered by the thesis/collection of published works and on the field of study in which the research lies. Where for reasons of sickness, disability or other valid cause, the Research Degrees Award Board is satisfied that a candidate would be under serious disadvantage if required to undergo a viva voce examination, it may approve an alternative form of examination. Such approval shall not be given on the grounds that the candidate's knowledge of the language in which the thesis is presented is inadequate.

Second stage (viva voce)

K16.6.3 The examiners shall not recommend that a candidate fail outright without holding a viva voce examination or other alternative examination.

K16.6.4 The examining panel will meet for a period of at least 30 minutes prior to the viva in order to plan the viva. The Chair is responsible for ensuring that the viva is conducted according to the Academic Regulations. Should any member of the examining panel have concerns about the academic integrity of the candidate's work which they have not previously raised in the preliminary report, they should raise this with the Independent Chair who will act in accordance with Academic Procedures at K17.

K16.6.5 A viva voce examination shall normally be held in English and on a campus of the University or the campus of an affiliated institution where the candidate has been registered for the award through that institution. The Director of Corporate and Academic Services/Academic Registrar may grant permission for a viva to be held elsewhere in the UK or abroad ("off campus") where there is good reason. Any decision to hold a viva voce examination off campus is subject to the appointment of an experienced internal examiner, the agreement of the candidate and all the examiners, and the appointment of an Independent Chair. Where it is proposed to hold a viva off campus the candidate and/or Director of Studies shall submit an application to the Officer to the Research Degrees Award Board in the first instance using the appropriate form.

K16.6.6 The Officer to the Research Degrees Award Board may, in exceptional circumstances, grant permission for one examiner to be available at a viva by video link, subject to the written agreement of both the candidate and the Independent Chair, and to the technology being of a satisfactory standard. In the event that the technology does not permit the viva to be conducted with the involvement of all parties to a satisfactory standard, the viva should be stopped and rearranged. It is not permissible for the candidate to be interviewed via video link or Skype.

K16.6.7 The supervisors and a representative of the Research Degrees Award Board may, with the consent of the candidate, attend the viva voce examination but may not participate in the discussion with the candidate unless at the invitation of the Chair. The representative of the Award Board shall remain whilst the examiners decide on their recommendation on the award but shall not participate in that discussion. The supervisors may not remain whilst the examiners decide on their recommendation on the award.

K16.6.8 Neither candidate nor supervisors may be present during the panel's deliberations. The conduct of the viva voce examination is at the discretion of the Independent Chair in

consultation with the examiners. Possible outcomes are as listed in K16.7.2R and in all cases the panel will complete and sign 'The Recommendation of the Examiners'.

K16.6.9 If the examiners agree on the outcome of the examination they shall, at its conclusion, submit on the designated form a joint report and recommendation relating to the award.

K16.6.10 The preliminary reports and the joint recommendation of the examiners shall together provide sufficiently detailed comments on the scope and quality of the work to enable the Research Degrees Award Board to be satisfied that the recommendation chosen is correct. The joint report shall be submitted to the Officer to the Research Degrees Award Board.

K16.6.11 If the examiners do not agree, they shall submit separate reports and recommendations on the designated forms. The Officer shall submit them to the Research Degrees Award Board for a decision in accordance with K16.7.3R.

K16.6.12 The Independent Chair must ensure that the examiners' report is duly completed and submitted to the Officer immediately after the examination.

2) Examples of doctoral assessment criteria

Birkbeck, University of London

Note: these criteria are taken directly from the QAA doctoral qualification descriptor, with the exception of the last bullet point, which has been extracted from the additional text supporting the descriptor (see Annex n below)

Doctorates are awarded for the creation and interpretation of knowledge, which extends the forefront of a discipline, usually through original research. Holders of doctorates will be able to conceptualise, design and implement projects for the generation of significant new knowledge and/or understanding.

A doctoral degree will be awarded to students who have demonstrated:

- The creation and interpretation of new knowledge, through original research, or other advanced scholarship, of a quality to satisfy peer review, extend the forefront of the discipline, and merit publication.
- A systematic acquisition and understanding of a substantial body of knowledge which is at the forefront of an academic discipline or area of professional practice.
- The general ability to conceptualise, design and implement a project for the generation of new knowledge, applications or understanding at the forefront of the discipline, and to adjust the project design in the light of unforeseen problems.
- A detailed understanding of applicable techniques for research and advanced academic enquiry.

Typically, holders of the qualification will be able to:

- Make informed judgments on complex issues in specialist fields, often in the absence of complete data, and be able to communicate their ideas and conclusions clearly and effectively to specialist and non-specialist audiences.
- Continue to undertake pure and/or applied research and development at an advanced level, contributing substantially to the development of new techniques, ideas, or approaches.

They will have:

- The qualities and transferable skills necessary for employment requiring the exercise of personal responsibility and largely autonomous initiative in complex and unpredictable situations, in professional or equivalent environments.
- *The qualities needed for employment requiring the ability to make informed judgments on complex issues in specialist fields, and innovation in tackling and solving problems.*

Durham University

By Thesis:

Candidates are required to demonstrate the ability to conduct original investigations, to test or explore ideas / hypotheses (whether their own or those of others), and to understand the relationship of the theme of their investigations to a wider field of knowledge. The thesis should include an original and significant contribution to knowledge, for example through the discovery of new knowledge, the connection of previously unrelated facts, the development of new theory, or a new analysis of older views. It should also include substantial matter worthy of publication, though it need not be submitted in a form suitable for publication.

University College London

The following is an extract from UCL's regulatory framework for the assessment and examination of Postgraduate Research students

Chapter 5: Assessment Framework: Research Programmes

Section 1: Research Degrees Assessment: PhD/EngD/MPhil/MD(Res)

1.1.2 Requirements of a PhD/EngD Thesis

1. A thesis for the awards of EngD or PhD degree shall be examined in accordance with the criteria prescribed by UCL and the thesis shall demonstrate that it:
 - a) is genuinely the work of the candidate;
 - b) shows a student's capacity to pursue original research in the field of study based on a good understanding of the research techniques and concepts appropriate to the discipline;
 - c) embodies the results of a research programme which may reasonably be expected of a student after three years of full-time study or the part-time equivalent, formulated and carried out by the student in consultation with the supervisors;
 - d) consists of a student's own account of their investigations, the greater proportion of which shall have been undertaken during the period of registration under supervision for the degree;
 - e) represents a distinct and significant contribution to the subject, whether through the discovery of new knowledge, the connection of previously unrelated facts, the development of new theory, or the revision of older views;
 - f) shows the exercise of critical judgement with regard to both a student's own work and that of other scholars in the field;
 - g) is an integrated whole and presents a coherent argument;
 - h) gives a critical assessment of the relevant literature, describes the method of research and its findings, includes discussion on those findings and indicates in what respects they appear to the student to advance the study of the subject; and, in so doing, demonstrates a deep and synoptic understanding of the field of study, (a student being able to place the thesis in a wider context), objectivity and the capacity for judgement in complex situations and autonomous work in that field.
 - i) is satisfactory in its literary and/or technical presentation and structure with a

- full bibliography and references;
- j) takes due account of previously published work on the subject;
- k) makes clear the sources from which information has been derived, the extent to which the work of others has been used, and the areas which are claimed as original;
- l) contains an element which might, after any necessary revision, merit publication in a medium appropriate to the discipline (for example as a monograph or as a number of articles in learned journals);
- m) shows a student's ability to design and implement an independent research project.

University of Reading

Note: the following is an extract from 'Examining PhDs and other research programme theses: Good Practice Guide'; February 2014

Assessing the thesis

According to the QAA, doctoral degrees are awarded to students who carry out original research, extend the forefront of their discipline and merit publication. In Reading, PhD candidates are expected to demonstrate each of the following:

- the creation and interpretation of new knowledge, through original research or other advanced scholarship, of a quality to satisfy peer review, to extend the forefront of their discipline, and to merit publication in an appropriate form
- a systematic acquisition and understanding of a substantial body of knowledge which is at the forefront of the discipline or area of professional practice
- the general ability to conceptualise, design and implement a project for the generation of new knowledge, applications or understanding at the forefront of the discipline, and the ability to adjust the project design in the light of unforeseen problems
- a comprehensive understanding of techniques applicable to their own research or advanced scholarship.

Probably the four key overall attributes for assessment are:

- contribution to knowledge
- originality and creativity
- integration and coherence presentation and
- clarity

The key things to be looking for when assessing the thesis are;

- is the problem worth addressing?
- is it being addressed in an appropriate way?
- is the literature review comprehensive and up to date, and does it show understanding of the theoretical context?
- does the candidate make explicit links between the review and his or her research question/s and study design?
- is there evidence of critical appraisal of the studies that are reviewed if there is empirical work, are the right methods being used, is the sample large enough, are the data presented clearly and analysed and interpreted appropriately, does the discussion illuminate the results, and do the stated conclusions follow from the results?

- if the work is theoretical, is the line of argument coherent, well expressed and does it develop logically?
- is the work ethical?
- is the work that of the candidate (i.e. not plagiarised)?

University of the West of England

Note: the University has developed qualification descriptors for all levels of awards that relate to the QAA qualification descriptors. The doctoral descriptor is set out below.

K3. Qualification descriptors

Definition: descriptors exemplify the outcomes and expectations of the main qualification at each level within Chapter A1 of the QAA Quality Code: The Framework for Higher Education Qualifications and demonstrate the nature of change between levels. They provide clear points of reference at each level and describe outcomes that cover the great majority of existing qualifications. The University's research degrees align to the descriptors for Masters (MPhil) and Doctoral level awards.

K3.2. Doctoral descriptor

K3.2.1 The award of a doctorate of the University (other than a Higher Doctorate) requires that a candidate should demonstrate that he / she:

- has conducted enquiry leading to the creation and interpretation of new knowledge through original research or other advanced scholarship, shown by satisfying scholarly review by accomplished and recognised scholars in the field;
- can demonstrate a critical understanding of the current state of knowledge in that field of theory and / or practice;
- shows the ability to conceptualise, design and implement a project for the generation of new knowledge at the forefront of the discipline or field of practice including the capacity to adjust the project design in the light of emergent issues and understandings;
- can demonstrate a critical understanding of the methodology of enquiry;
- has developed independent judgement of issues and ideas in the field of research and / or practice and is able to communicate and justify that judgement to appropriate audiences and,
- can critically reflect on his / her work and evaluate its strengths and weaknesses including understanding validation procedures.

3 Examination outcomes and recommendations

Oxford University

General Regulations for the Degree of Doctor of Philosophy (extract)

§7. Examination of Students for the Degree of Doctor of Philosophy

Having completed the examination of a candidate for the first time, the examiners may make any one of recommendations (i), (ii), or (iv) below only. Having completed the examination of a candidate who has revised and re-submitted his or her thesis, the examiners may make any one of recommendations (i)-(vi). The recommendations are:

(i) that the board should grant the candidate leave to supplicate for the Degree of Doctor of Philosophy. In making this recommendation, the examiners must include in their report statements that:

1. the student possesses a good general knowledge of the particular field of learning within which the subject of the thesis falls;
2. the student has made a significant and substantial contribution in the particular field of learning within which the subject of the thesis falls;
3. the thesis is presented in a lucid and scholarly manner;
4. in their opinion the thesis merits the Degree of Doctor of Philosophy;
5. the student has presented a satisfactory abstract of the thesis.

Examiners shall bear in mind that their judgement of the substantial significance of the work should take into account what may reasonably be expected of a capable and diligent student after three or at most four years of full-time study in the case of a full-time student, or eight years in the case of a part-time student.

(i) (a) **Minor corrections** If the examiners are satisfied that the candidate's thesis is of sufficient merit to qualify for the degree but consider, nevertheless, that before the thesis is deposited the candidate should make minor corrections (which are not sufficiently substantial to justify reference back for re-examination and which should be capable of completion within one month), they must require the candidate to correct the thesis to their satisfaction before they submit their report. If the candidate has not completed these corrections within one calendar month of the date of receipt of the list of minor corrections from the examiners, his or her name shall be removed by the Registrar from the Register of Students for the Degree of Doctor of Philosophy, provided that the board may, on good cause shown by the candidate, grant an extension of time of one further calendar month in which the candidate may fulfil this requirement before the removal of his or her name from the Register. No subsequent extension shall be granted, but it shall be open to a candidate who has failed to fulfil this requirement within those one or two months in total, as the case may be, to apply to the board for reinstatement as a Student for the Degree of Doctor of Philosophy, with the support of his or her society and supervisor, upon submission to the Registrar of a copy of his or her thesis incorporating the required corrections, and upon payment of such reinstatement fee as may from time to time be prescribed by Council by decree. Permission to supplicate shall not be granted until this fee has been paid;

(i) (b) **Major corrections** If the examiners are satisfied that the candidate's thesis is of sufficient potential merit to qualify for the degree but consider, nevertheless, that before the thesis is deposited the candidate should make major corrections (which are not sufficiently substantial to justify reference back for re-examination and which should be capable of completion within six months), they should report this preliminary recommendation to the board with a description of the major corrections which they require the candidate to make before they confirm their recommendation. Where the examiners make this recommendation, and the board, considering the extent and nature of the major corrections, takes the view that the recommendation ought to be reference of the thesis back to the candidate in order that he or she may

revise it for re-examination, the board may, exceptionally, ask the examiners to review their recommendation. If the candidate has not completed these corrections within six calendar months of the date of receipt of the list of major corrections from the examiners, his or her name shall be removed by the Registrar from the Register of Students for the Degree of Doctor of Philosophy, provided that the board may, on good cause shown by the candidate, grant an extension of time of up to three further calendar months in which the candidate may fulfil this requirement before the removal of his or her name from the Register. No subsequent extension shall be granted, but it shall be open to a candidate who has failed to fulfil this requirement within those six or nine months in total, to apply to the board for reinstatement as a Student for the Degree of Doctor of Philosophy, with the support of his or her society and supervisor, upon submission to the Registrar of a copy of his or her thesis incorporating the required corrections, and upon payment of such reinstatement fee as may from time to time be prescribed by Council by decree. Permission to supplicate shall not be granted until this fee has been paid. Where a recommendation of approval subject to major corrections has been made, the examiners, on receipt of the corrected thesis on the first occasion, may conclude and report one of the following:

- (a) the original recommendation is now fully substantiated;
- (b) the work as submitted still requires minor corrections prior to confirmation of the original recommendation and a further one month may be allowed for this from that date;
- (c) the work as now submitted still requires major corrections prior to confirmation of the original recommendation and a further six months may be allowed for this from that date.

Recommendation (a) or (b) may be made without a further oral examination on condition that both examiners have reviewed and approved the major corrections. A further oral examination must be held if either of the examiners requires it or if the recommendation is likely to be (c). In the exceptional instance where (c) applied and the outcome of the second and final submission of corrections still remained unsatisfactory, the examiners will amend their original recommendation to (ii) below.

(ii) that the board should offer the candidate a choice between (a) reference of the thesis back to him or her in order that he or she may revise it for re-examination for the Degree of Doctor of Philosophy, and (b) leave to supplicate for the Degree of Master of Letters or of Master of Science, as appropriate, on the basis that the thesis has not reached the standard required for the Degree of Doctor of Philosophy but has nevertheless reached that required for the Degree of Master of Letters or of Master of Science.

- (a) If the board adopts this recommendation, and the student chooses to revise the thesis for re-examination for the Degree of Doctor of Philosophy, the student shall retain the status and obligations of a Student for the Degree of Doctor of Philosophy and shall be permitted to apply again for the appointment of examiners, in accordance with the procedure laid down in this sub-section, not later than the sixth term after that in which the board gave permission so to reapply. If such permission shall have been given by a board during a vacation, it shall be deemed to have been given in the term preceding that

vacation. Accompanying the revised thesis at re-submission should be a separate report indicating the specific changes made. For students in the Humanities, Medical Sciences and Social Sciences Divisions and the Department for Continuing Education, the word limit for the accompanying report shall be 1,000 words; for students in the Mathematical, Physical and Life Sciences Division, the word limit shall be 2,000 words;

(b) If the board adopts this recommendation and the student chooses leave to supplicate for the Degree of Master of Letters or Master of Science by Research, the examiners may still determine that before the thesis is deposited the candidate should make minor corrections in accordance with the regulations under (i) above.

(iii) that the board should refer the student's thesis back in order that he or she may present it for re-examination for the Degree of Master of Letters or of Master of Science, as determined by the examiners (if appropriate), only. If the board adopts the recommendation the student shall be transferred forthwith to the status of Student for the Degree of Master of Letters or Student for the Degree of Master of Science as the case may be, and shall be permitted to apply for permission to supplicate for the Degree of Master of Letters or Master of Science in accordance with the provisions of the appropriate regulation. If such permission shall have been given by a board during a vacation, it shall be deemed to have been given in the term preceding that vacation. The word limit for a thesis resubmitted under this provision shall be that specified by the DPhil regulations under which it was originally submitted;

(iv) that the board should refer the student's thesis back in order that he or she may present it for re-examination either under (ii) above for the Degree of Doctor of Philosophy or, if the student chooses, under (iii) above for the Degree of Master of Letters or of Master of Science only. The board shall adopt such a recommendation only if it is fully satisfied that the thesis as it stands is not of the standard required for the Degree of Doctor of Philosophy, nor for the Degree of Master of Letters or of Master of Science as the case may be, but that the candidate could reach the standard required for the Degree of Doctor of Philosophy. If such permission shall have been given by a board during a vacation, it shall be deemed to have been given in the term preceding that vacation;

(v) that the thesis has not reached the standard required for the Degree of Doctor of Philosophy but has nevertheless reached that required of the Degree of Master of Letters or of Master of Science, and that the candidate may be granted leave to supplicate for one of the latter degrees on the basis of the thesis as it stands; the examiners may still determine that before the thesis is deposited the candidate should make minor corrections in accordance with the regulations under (i) above.

(vi) that the student's application for leave to supplicate should be refused.

7. If the examiners recommend reference back of the student's application under clause 6 (ii) or (iii) or (iv) above, they shall annex to their report to the board a statement (for transmission to the candidate) setting out the respects in which the thesis falls below the standard required for the degree in question, and what changes are necessary for it to reach that standard, save that examiners of a thesis submitted for the first time may, in exceptional circumstances, and

notwithstanding a recommendation under clause 6 (ii) or (iv) above, certify that they are unable to indicate how the thesis might be changed, within the time allowed, in order to reach the required standard for the degree of Doctor of Philosophy.

8. On receipt of the examiners' report the board shall reach a decision on whether to accept the examiners' recommendation, provided that no candidate shall be given leave to supplicate for the Degree of Doctor of Philosophy unless the examiners have made the statements required in clause 6 (i) above.

9. A candidate who has been granted leave to supplicate by a board shall be required to submit to the Examination Schools a copy of his or her thesis, incorporating any amendments or corrections required by the examiners and approved by the board, with a view to deposit in the Bodleian or other appropriate university library. [For candidates admitted on or after 1 October 2007: candidates are also required to submit an electronic copy of their thesis to the Oxford Research Archive, unless an exception to this requirement has been granted by the Proctors.] [For candidates supplicating on or after 1 July 2013: candidates are also required to submit the library copy to the Examination Schools and where applicable the electronic copy of the thesis to the Oxford Research Archive no later than the end of the fifth day before the date of the degree ceremony booked by the candidate for conferral of their degree.] Permission to supplicate shall in all cases be conditional upon fulfilment of these requirements.

10. In an exceptional case in which a board is unable to accept the examiners' recommendation, or in which the examiners cannot reach an agreed recommendation, the board shall have power to appoint one or two new examiners as it deems necessary, to conduct such further examination of the candidate as the board may require. The board shall make a report on any such case to the Education Committee.

11. The board may exempt a candidate who is being re-examined under the provisions of clause 6 (ii)–(v) above from the oral examination, provided that the examiners are able to certify that they are satisfied without examining the candidate orally that they can recommend to the board in the terms required by clause 6 (i) above that he or she be given leave to supplicate for the Degree of Doctor of Philosophy.

12. It shall be the duty of the Registrar to notify the candidate of the board's decision as soon as may be. The Registrar shall also be responsible for publishing at the end of each academic year the names of those candidates to whom permission to supplicate has been granted during that year, together with a statement of the subject of the thesis written by each.

13. When, on the conclusion of the investigation of a complaint made by a candidate, the Proctors recommend that a candidate be re-examined, the board shall have power to hold a new examination.

¹ Here and hereafter in these regulations, in the case of Fine Art candidates offering studio practice as part of their submission, and in the case of Music candidates offering a portfolio of musical compositions as part of their submission, 'thesis' shall be understood to include the totality of the candidate's submission.

Manchester University

Examination of Doctoral Degrees Policy June 2016 (extract)

PART FIVE: EXAMINER RECOMMENDATIONS

22. Recommendations for doctoral degrees – first examination

There are three categories of recommendations for doctoral degrees: 'A' (award), 'B' (refer) and 'C' (reject). Within each category, examiners must select a sub-recommendation, as follows.

22.1 CATEGORY A: AWARD (recommendation A(i) and A(ii))

22.1.1 Award with no corrections (recommendation A(i))

The examiners should select recommendation A(i) if the thesis is satisfactory in every way and there are no revisions to be made to it.

The examiners may recommend the award if they are satisfied that the thesis is satisfactory in every way and that:

- the candidate possesses an appropriate knowledge of the particular field of learning within which the subject of the thesis falls;
- the research which is reported in the thesis contributes a substantial addition to knowledge;
- the results of the research show evidence of originality and independent critical judgement;
- the thesis is presented in a lucid and scholarly manner;
- the thesis has been submitted in the form prescribed by University regulations and policy;
- Any contribution of others has been clearly detailed;
- no part of the thesis has previously been submitted for the award of a degree at this or any other University (students should refer to the Publication Format – Guiding Principles for further information);
- the thesis and the work reported in it are the candidate's own.

Following receipt of an A(i) award recommendation the final version of the thesis should be submitted by the student via eScholar, normally within 10 working days.

22.1.2 Award subject to minor corrections (recommendation A(ii))

Recommendation A(ii) should be selected if the examiners are satisfied that the thesis meets the criteria for the degree (see section 22.1.1 above) but some minor corrections are necessary to the thesis. The corrections, in the view of the examiners, and taking into account the guidance given below, should not be sufficiently serious to merit a recommendation for resubmission and re-examination under Category B.

Minor corrections permissible under box A(ii) include:

- typographical errors; however, if the errors, although trivial individually, are so numerous as to suggest carelessness on the part of the candidate or so intrusive as to distract the reader's attention from the argument of the thesis, the examiners would be fully justified in making a recommendation under Category B rather than box A(ii);
- minor amendments and/or replacement of, or additions to, the text or to references or diagrams;
- other more extensive corrections may be made as long as they do not require significant (as defined by the examiners) re-working or re-interpretation of the intellectual content of the thesis. If more substantial revisions are required, the examiners should select one of the recommendations under Category B.

A list of corrections must be provided by the examiners in section four of the Examiners' Report Form for the benefit of the candidate. Once carried out by the candidate, the corrections must be approved by the internal examiner without the need for a further oral examination.

The time permitted for minor corrections to be completed by the candidate is normally no more than four weeks from the date the candidate receives the list of corrections in the University's progression monitoring system. In exceptional circumstances, where there are more than four weeks of work required of the student the candidate may be given 12 weeks to complete the revisions. The examiners' decision to allow 12 weeks for minor corrections to be completed, must be based on the quantity of the work required and length of time of which it is feasible to complete the corrections.

The Graduate Office must receive notification that the minor corrections have been approved by the internal examiner via the University's progression monitoring system within 2 weeks of the candidate submitting the revisions.

In examinations where there is no internal examiner, the independent chair must ensure that minor revisions are approved and confirmed in the University's progression monitoring system by an external examiner or other person nominated by the School.

The examiners' decision to recommend an A(ii) should be made on the grounds that the thesis will NOT require a further examination. The decision whether to recommend an A(ii) as opposed to a B(i) must not be determined by the candidate's personal circumstances.

The candidate is expected to be available in the period after the oral examination to complete minor revisions as part of their responsibilities in the examination of their degree. In very exceptional circumstances, the candidate may apply to the appropriate School or Faculty Graduate Office for permission to submit the corrected thesis later than the four-week or 12 week deadline given.

Candidates and examiners should refer to the appropriate School or Faculty Graduate Office if further guidance on minor corrections is required (see appendix one for contact details).

22.2 CATEGORY B: REFER FOR RE-EXAMINATION (recommendations B(i), B(ii) and B(iii))

Referral under recommendation B requires the candidate to resubmit the thesis for re-examination. A candidate will be permitted to resubmit on only one occasion. See the University's Resubmission and Re-examination of Postgraduate Research Degrees Policy for details of resubmission and re-examination.

Examiners are required to make one of the following recommendations under category B:

- B (i) that the thesis is satisfactory in substance, but defective in presentation and/or content and does not require a further oral examination;
- B (ii) that the thesis is satisfactory in substance, but defective in presentation and/or content and requires a further oral examination;
- B (iii) that the thesis is unsatisfactory in substance, defective in presentation and/or content and requires further research and a further oral examination.

If examiners recommend that the candidate will require a further oral examination upon resubmission of the thesis (recommendations B(ii) and B(iii)), the examiners may later, if in joint agreement and if their recommendation is to award the degree, dispense with the oral examination after assessment of the resubmitted thesis. Examiners must inform the appropriate School or Faculty Graduate Office staff as soon as possible if they wish to dispense with the oral examination so that the candidate can be informed and any refund given.

For category B recommendations, examiners must submit a statement as part of the joint Examiners' Report form (see section 23) detailing the required revisions. Examiners should offer guidance to the student when their recommended changes will impact on the word limit of the thesis; this may include suggestions on which parts of the thesis can be reduced to allow for any additions.

For recommendations B(i) and B(ii), the candidate is normally required to revise and resubmit the thesis for the doctoral degree within six months of receiving the examiners' statement detailing the required revisions (see section 23). Examiners may make a recommendation to extend that period so that the candidate is required to revise and resubmit the thesis for the doctoral degree within 12 months of receiving the examiners' statement detailing the required revisions from the appropriate School or Faculty Graduate Office. The decision to exceptionally allow additional time for the student to revise and resubmit is made by the appropriate committee and must be based on the quantity of the work required and length of time it is feasible to complete the work.

Any student who wishes to request an extension to their resubmission deadline, due to mitigating circumstances, should make the request via the normal extension application process within the School.

For recommendation B(iii), the candidate is required to revise and resubmit the thesis for the doctoral degree within 12 months of receiving the examiners' statement detailing the required revisions (see section 23) from the appropriate School or Faculty Graduate Office.

22.3 CATEGORY C: REJECT (recommendations C(i), C(ii), C(iii) and C(iv))

Where examiners are not satisfied that the thesis and oral examination have met the standards required, and have not found evidence that the thesis could be corrected under category A or B, they may recommend:

- C(i): Reject but award the degree of Master of Philosophy (MPhil). For recommendation C(i), examiners must refer to the degree criteria in the MPhil Ordinances and Regulations and justify their decision to recommend the award of MPhil. The candidate must resubmit the thesis, with a new title page, indicating that it is a master's (and not a doctoral) thesis.
- C(ii): reject but award the degree of Master of Philosophy (MPhil) subject to minor corrections being made to the thesis. For recommendation C(ii), as part of the Examiners' Report Form examiners must include a statement detailing the required corrections (see section 23) Examiners must refer to the degree criteria in the MPhil Ordinances and Regulations and justify their decision to recommend the award of MPhil, subject to minor corrections.

The time permitted for minor corrections to be completed by the candidate is normally no more than four weeks from the date the candidate receives the list of revisions. Where there is more than four weeks of work required of the student the candidate may be given 12 weeks to complete the corrections. The notification that the minor corrections have been approved by the internal examiner must be submitted to the University's progression monitoring system within this timeframe.

The decision to allow 12 weeks for minor corrections to be completed, must be based on the quantity of the work required and length of time it is feasible to complete the corrections.

On completion and approval of minor corrections, the candidate must upload the thesis to eScholar, with a new title page, indicating that it is a master's (and not a doctoral) thesis.

- C(iii): reject but invite the candidate to revise and resubmit the thesis for examination for the degree of Master of Philosophy (MPhil). For recommendation C(iii), examiners must submit a statement detailing the required revisions (see section 23) as part of the Examiners' Report Form. The candidate is required to revise and resubmit the thesis for examination for the MPhil degree within six months of receiving the examiners' statement via the University's progression monitoring system detailing the required revisions, and, where appropriate, attend one further oral examination. The candidate will be permitted to resubmit on only one occasion.
- C(iv): reject and no resubmission be permitted. For recommendation C(iv), examiners must justify their decision not to recommend the doctorate in the Examiners' Report Form. When the recommendation of the examiners is not to award a doctoral degree, the internal examiner or independent chair where appropriate will normally be invited to the next meeting of the appropriate School or Faculty postgraduate research degrees committee to assist in its consideration of the case and to answer any questions.

23. Examiners' statement: only applicable for category B recommendations and recommendations C(ii) and C (iii)

Where examiners recommend any of the category B or the category C recommendations C(ii) or C(iii), examiners must produce a written statement of required revisions. This will then be provided to the candidate via the University's progression monitoring system. Examiners should give sufficient detail of the defects of the original submission and recommend ways in which the thesis should be corrected in order to make a satisfactory revision of the thesis. The statement must be suitable to form the basis of the subsequent re-examination.

The statement should specify to the candidate changes that need to be made to the thesis before resubmission, although the statement does not need to descend to the level of specifying every revision to spelling, grammar, etc, where these are numerous, and can state requirements in general terms, where appropriate. The statement should be in a form suitable for communication to the candidate.

The examiners must agree this statement of required revision and ensure it is included as part of the joint examiners report, within five working days of the examination. The candidate can view the report in the University's progression monitoring system with the official written notice of the recommendation once it has been confirmed by the appropriate School or Faculty postgraduate research degrees committee

University of the West of England

Academic Regulations and Procedures 2016/17 (extract)

K16.7 Examiners recommendations

K16.7.1R The Research Degrees Award Board will make recommendations for the award of research degrees to Academic Board on the basis of the reports and recommendation of the examiners following the viva voce examination and in consideration of outcomes from taught components.

K16.7.2R Following the viva voce examination the examiners may recommend that:

A The candidate fulfils the criteria for the award on which they are registered: Pass; including corrections; amendments - minor and major outcomes

The candidate fulfils the Doctoral/MPhil award criteria and examiners may recommend that the candidate be awarded the degree:

- i. Without further correction or amendment;
- ii. Subject to **correction** of presentational/typographical errors within the material* (maximum 4 weeks FT/6 weeks PT). Corrections to be approved by one or both/all examiners;
- iii. Subject to **minor amendment** of the material* as indicated by the examiners and which can reasonably be completed within a maximum 12 weeks FT / 18 weeks PT. Amendments to be approved by one or both / all examiners;
- iv. Subject to **major amendment**. The material* submitted displays some deficiencies of content , analysis and / or presentation in areas specified by the examiners requiring additional work which can reasonably be expected to be completed within a maximum 6 months FT / 9 months PT. No re-examination is required, amendments to be approved by all examiners.

B. The candidate does not currently fulfil the criteria for the award on which they are registered:

Referred for resubmission and re-examination

The candidate does not currently fulfil the Doctoral/MPhil criteria and the material* as submitted displays significant deficiencies of content and/or presentation in areas specified by the examiners. The candidate is permitted to revise and re-submit the material for the degree and be re-examined on one further occasion with or without viva. Revisions indicated by examiners may reasonably be expected to be completed within a maximum 12 months FT/18 months PT. The re-examination shall be of the submitted material as a whole and by all examiners.

C. Additional outcomes for PhD or DPhil examination only:

i. MPhil with amendments

The candidate does not fulfil the doctoral award descriptor criteria but does meet the criteria for MPhil and may be recommended for this award subject to amendment of the material* in a manner and to a timescale as recommended by the examiners (**up to** a maximum of 6 months FT/9 months PT). No further examination is required. Amendments to be approved by one or both/all examiners;

Or

ii. Resubmit and be examined for MPhil

The candidate does not fulfil the doctoral award criteria but has the potential to meet the criteria for MPhil and may revise and resubmit the material* as indicated by the examiners for examination for the award of MPhil (within a maximum of 12 months FT/18 months PT). The examination shall be of the submitted material as a whole, shall include a viva and shall be by all examiners.

D. Degree not awarded

The candidate is not awarded the degree and is not permitted to be re-examined. (Unsuccessful candidates for DPhil/MPhil by publication may be permitted to re-apply after a period of three years.)

*NB. Material in the case of DPhil/MPhil by publication material refers to the critical commentary element of the submission.

Disagreement between examiners following a first viva examination

K16.7.3R Where the examiners' recommendations are not unanimous, this shall be reported by the Independent Chair to the Research Degrees Award Board (RDAB) which may:

- a. uphold a majority recommendation (provided that the majority includes at least one external examiner);
- b. uphold the recommendation of the external examiner; or
- c. appoint an independent external assessor who shall review the thesis and make an independent report together with an outcome recommendation to RDAB. No further viva voce examination of the candidate shall take place within the first attempt. RDAB will consider all reports and agree an outcome decision in accordance with regulations at K16.7.2R.

Application of the regulations

K16.7.4 Where the candidate is awarded the degree subject to correction or amendment (K16.7.2R.A ii. – iv.), or is referred for re-submission and re-examination (K16.7.2R B), or receives an outcome recommendation (K16.7.2R C i. or ii), the Independent Chair will be responsible for the co-ordination of type-written feedback for the candidate using the electronic template supplied detailing the requirements of the examining panel as to alterations and additional work, together with any other guidance that examiners wish to pass on to the candidate as appropriate. All examiners must agree the content of the written feedback.

K16.7.5 For the sake of convenience annotations noted on an examiner's copy of the thesis indicating minor presentational/typographical errors only may be passed to the candidate, but this will be sent concurrent with the written feedback by the Officer of the Research Degrees Award Board to whom the annotated copy must be forwarded by the Independent Chair. An annotated thesis can only be accepted in this way in conjunction with the written feedback, not as a substitute for it.

K16.7.6 Where examiners wish any specific comments contained within their pre-viva preliminary report to be made available to the candidate in post-viva feedback they must indicate this clearly to the Independent Chair. It will not otherwise be made available to the candidate.

K16.7.7 The written feedback will be communicated to the candidate and the Director of Studies by the Officer to the Research Degrees Award Board.

K16.7.8 Deadlines for the submission of required corrections, amendments or resubmission of material as a whole shall be calculated from the date of notification of the outcome of the viva examination by the Research Degrees Award Board.

K16.7.9 Amendments should be made in accordance with specific requirements of the examiners. However, the candidate remains ultimately responsible for deciding the manner in which to improve the material and when these shall be submitted within the maximum time stipulated at K16.7.2R. Candidates may choose to submit before the maximum time allowed.

Minor amendment

K16.7.10 Where the award outcome is subject to minor amendment (K16.7.2R.A.iii.), the candidate meets the award criteria and amendment shall normally be restricted to amendment of representational errors or re-presenting and restructuring existing text only; new work should not normally be required.

Major amendment

K16.7.11 Where an award outcome is subject to major amendment (K16.7.2R.A.iv.) the candidate meets the award criteria but the material submitted does not reflect the quality of the research undertaken and requires some additional work. This outcome may include a measure of new work including limited extra research or analysis, undertaking some new experiments or repeating existing ones, or re-writing sections of material. This, however, should not amount to a significant extension or fundamental change in direction of the original research, and will not involve a complete re-write of the material as a whole.

K16.7.12 Where the candidate is awarded the degree subject to major amendment, the Research Degrees Award Board may recommend that some further supervision may be beneficial. Where the candidate chooses to take up this recommendation they will be required to pay the appropriate pro-rata fee.

Unsatisfactory amendments

K16.7.13 Should the amendments submitted, whether major or minor, not be deemed satisfactory by the examiners the candidate will be given a further 12 weeks in order to make the necessary adjustments. If after this additional time, the examiners confirm that they are not content with the amended material the candidate will be deemed not to have complied with the requirements of the award and may be required to withdraw by the Research Degrees Award Board and the degree not be awarded.

Referred for resubmission and re-examination

K16.7.14 Where the award outcome is referred for resubmission and re-examination (K16.7.2R.B.), the candidate does not currently meet the relevant award criteria but with substantial new work and re-writing of the material the examiners judge that they have the potential to do so at re-examination. Revision of the material may relate to theoretical and/or methodological aspects and new work may include any or all of the following: new research and/or new data, fieldwork or practice, new analysis, substantial new literature. Exceptionally the examiners may recommend that the candidate be re-examined by viva only, without having to resubmit revised material.

K16.7.15 A candidate who is referred for resubmission and re-examination with or without a further viva shall only be exempt from the additional viva where the examiners unanimously agree, after re-examining the resubmitted thesis, that there is no need for it, and that the degree can be awarded. However, a candidate may not be failed outright at resubmission and the degree not awarded (outcome K16.7.2R.D.) without the opportunity to undergo a further viva.

K16.7.16 Where a resubmission viva is required, the Director of Studies may request a copy of the examiners' resubmission preliminary reports from the Graduate School to assist the candidate's preparation for the resubmission viva.

K16.8 Re-examination of thesis or collection of published works

K16.8.1R A candidate who submits a thesis or collection of published works for re-examination shall be required to pay the appropriate fee.

K16.8.2R One re-examination may be permitted by the Research Degrees Award Board subject to the candidate submitting for re-examination within 12 months FT/18 months PT from the date of the latest part of the first examination.

K16.8.3R The Research Degrees Award Board may require that an additional external examiner be appointed for the re-examination.

K16.8.4R The Research Degrees Award Board may, where there are good reasons, approve an extension of the resubmission period.

K16.8.5R In all other respects the re-examination shall be conducted as the first examination, excepting where the examiners' recommendations following a resubmission viva are not unanimous.

K16.8.6R Following the completion of the re-examination the examiners may recommend in accordance with the provisions set out in K16.7.2R, excepting that outcome option B shall not apply.

Disagreement between examiners following a resubmission viva

K16.8.7R Where, following a resubmission viva, the examiners' recommendations are not unanimous, the Research Degrees Award Board (RDAB) may:

- a. uphold a majority recommendation (provided that the majority includes at least one external examiner); or
- b. uphold the recommendation of the external examiner; or
- c. appoint an independent external assessor who shall review the resubmitted thesis and make an independent report together with an outcome recommendation to RDAB. No further viva examination of the candidate shall take place. RDAB will consider all reports and agree a final outcome as at K16.7.2R (save that option B shall not apply).

K16.8.8R Where the degree is not awarded, the examiners shall prepare an agreed statement of the deficiencies of the thesis or collection of published work and the reason for their recommendation, to be forwarded to the candidate by the Officer to the Research Degrees Award Board.