The Experience of Examining the PhD

An International Comparative Study of Processes and Standards of Doctoral Examination

Edited by Michael Byram and Maria Stoicheva

First published 2023

ISBN: 978-1-032-15697-2 (hbk)
ISBN: 978-1-032-16436-6 (pbk)
ISBN: 978-1-003-24856-9 (ebk)

Chapter 1

Introduction

MICHAEL BYRAM AND MARIA STOICHEVA

DOI: 10.4324/9781003248569-1

The Open Access version of chapter 1 was funded by Sofia University St. Kliment Ohridski.

This study is financed by the European Union-NextGenerationEU, through the National Recovery and Resilience Plan of the Republic of Bulgaria, project No BGRRP-2.004-000, coordinated by Sofia University St Kliment Ohridski.

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

1 Introduction

Michael Byram and Maria Stoicheva

The PhD is usually the highest academic award offered by universities. We say here 'usually' because, as this book will show, there are few if any generalisations about doctoral studies which stand the test of empirical investigation. We found that to be the case already in our previous book, *The Doctorate as Experience in Europe and Beyond* (Byram & Stoicheva, 2020), which concentrated on 'supervision, languages, identities', and here we focus on assessment, with a similar result. Yet if doctoral studies are to have some kind of international validity and a PhD is to be used as an international 'currency', in these times of globalisation and internationalisation, it would appear self-evident that some kind of generalisation about how doctoral studies are assessed is important.

It is not our aim here to propose an international system of assessment or even to make a recommendation that such a system should be constructed. Our aim is to provide a basis for discussing such matters founded on the actual experience of people in a wide range of universities and countries. For it seems to us that, although theoretical, conceptual analysis of assessment for doctoral studies is important, the experience and reflections of those involved are equally so.

This book thus provides 13 case studies of universities in 13 countries. Each case study, it is important to emphasise, is about one university, and doctoral studies in the humanities and social sciences – for instance departments of political science, education, literatures and languages, sociology – within that university. We are not offering case studies of whole countries, although we do ask our interviewees to talk about their experience in other universities in their own country or abroad. A different and complementary approach has been taken by Kumar et al. (2023).

In this introduction we shall, first, review previous research, both conceptual and empirical, on the assessment of doctoral studies and, in particular, the final 'product', i.e., the 'dissertation' or 'thesis'; the term varies from country to country and we shall use both. We shall then explain the research methodology including the selection of cases, the approach taken to data collection, and the process of data analysis and presentation. The findings of the project are then presented in two kinds of chapters: four transversal, comparative chapters dealing with specific themes and drawing on the case studies, and the 13 case studies. The case studies were written to follow a general pattern of presenting

DOI: 10.4324/9781003248569-1

the university, the regulatory framework as context and then the analysis of interviews, but each follows its own internal necessities.

Analysis of previous research[1]

Most of the literature on the PhD is to be found in empirical studies within the social sciences. One notable exception is a Crosscourt book-length study of the western 'PhD imaginary' (Kelly, 2017) which engages with 'works of the imagination' to gain understanding of how the PhD is conceptualised (p. 118) in the university and beyond. Dardy et al. (2002), too, refer to imaginative literature and cinema as a way of understanding the PhD experience. Neither Kelly nor Dardy et al., however, include the experience of the examination in this approach.

It is in the empirical studies, usually reported in journal articles, that we find a focus on examination, from submission through to communicating results (including delays in giving results Irvine, 2012). There is a shift in emphasis over time from analysing examiners' reports, through analysing the processes of the 'oral examination', and the characteristics of examiners, to analysing feedback from examiners for corrections, with a glance at the role of convenors/chairs of examining panels (Kumar et al., 2020). A frequently mentioned concern is to 'demystify' and make the 'secret' process, particularly the examination itself, 'more transparent' (Carter, 2008; Carter, B. & Whittaker, K. 2009; Dally et al., 2004; Golding et al., 2014; Morley, 2004; Murray, 2003; Tinkler & Jackson, 2000; Watts, 2012). There is also some concentration on 'QA' (quality assurance), but surprisingly little despite this having dominated bachelor's and master's courses in Anglo Saxon universities, and despite most of the research on assessment of the PhD having been done in Britain, Australia, and New Zealand, where QA is in general a major issue.

Part of the explanation for the different foci and methods is the lack of access to the oral examination, which is called 'the viva' in the Anglo-Saxon world, and 'the defence'[2] in many other countries. Although the defence which is usually 'public' – either in the sense of open to other members of the university or to them and members of the general public – the viva takes place in private and remains confidential. This means research methods are limited to text and discourse analysis, and analysis of perspectives and opinions of participants, using both qualitative and quantitative data, and do not include observational studies, although there is the occasional analysis based on notes taken during a viva (Sikes, 2017; Trafford, 2003). Since access to the defence is usually public, there is a potential for observational studies, although none have been found. Such studies would open a new dimension of research rather than filling the 'gap' in Anglo-Saxon studies because the nature of the event is quite different.

Hodgson (2020) is the only review article and is a 'synthesis' of research on examiner expectations with respect to the dissertation/thesis (but not the oral examination). Hodgson is, like a good number of other authors in this field, Australian, and in Australia, there is no viva. He develops the review in order to

develop what he calls a 'tool' to help in supervision. There is here, as in many other articles, the aim to 'demystify' the examining process, although he does not use this word. The analysis is 'based on a review of guidelines on PhD standards' and its purpose is to identify the 'kinds of things that examiners look for and expect in a PhD thesis' (p. 52). The 'tool' consists of a table of characteristics of a dissertation/thesis examiners look for, based on seven main themes:

- Mastery/command – comprehensive and exhaustive coverage and understanding of the subject;
- Argument – an overarching point, which is explained and defended;
- Coherence – consistency and clear links between parts and the whole;
- Independence – originality and ownership of the work;
- Criticality – evaluation of existing knowledge and self-evaluation of own work;
- Depth/breadth – thoroughness and completeness of work with sufficient detail;
- Clarity/accuracy – thesis is correctly written and sources accurately cited.

(ibid. p. 58)

Much of the article is in fact devoted to discussing supervision and how the tool might be used.

Apart from this review, other publications tend to focus on one theme and we shall analyse here, thematically. Such an approach – by theme irrespective of date – has advantages of clarity, although it is important to remember that the 'PhD' (and recent variations such as DBA, EdD, etc. (Hoddell et al. 2002) is changing and an a-chronological approach has disadvantages in this respect.

Comparisons and consistency in examining

Studies of comparability and consistency in examining processes are usually focused on a country and ask questions about practices within national boundaries. Holbrook and her colleagues carried out a large-scale analysis in Australia which is reported in several publications. Their approach is to combine qualitative analysis of examiners' reports with quantification of the categories and codes used, in order to establish tendencies within the corpus of 2121 examiners' reports on 804 theses across all broad fields of study. In general, Holbrook and colleagues found consistency and strong similarities in the various dimensions of examination. In Holbrook et al. (2008), the conclusion reached was that there was 'an extremely high level of consistency in examiner recommendations' (p. 45), but they also argued that there are problems and 'worst-case scenarios' which suggest a need for 'a clear code of practice and transparent procedures'. Furthermore, they suggested that, since examiners' comments and criteria were similar and generic for different disciplines, there should be no substantial problems in assessing interdisciplinary theses. In Holbrook et al. (2004), the research team established from their inductive analysis that there are agreed if

tacit criteria present in examiners' reports, and that 'there is no reason to write standards from first principles or impose them from without, because they can be reliably identified from what examiners say in their extensive reports' (p. 46). We shall return to the question of standards and benchmarks below.

In studies in Britain, comparisons of doctorate examining deal with the procedures, including for example appointment of examiners, role of supervisors in examining, and form of the oral examination/viva. Johnston (1997) analyses and comoares examiners' report. With the purpose of shedding light on the examining process, 'this mysterious but crucial component of the British doctorate' (p. 168), Tinkler and Jackson (2000) analyse documents from 20 universities and, broadly speaking, find considerable consistency. However, with closer consideration of details, they find much variation within the broad similarities. With respect to criteria for assessment, for example, all 20 universities required 'originality' in the thesis but only 6 required that there should be 'publishable' material. Another consistent feature is insistence on the 'independence' of 'external' examiners, i.e., examiners from other universities, but there is much variation in how independence is described and defined. Variation in the way the viva should be conducted is also evident, and the authors point out that, despite a widely held view that British university oral examinations are held in private, there are some universities where, in policy at least, the viva is 'public'. The interpretation of this term is however different from that in continental European universities and limits attendance to academic persons.

Tinkler and Jackson's (2000) conclusion that there is much variation in British universities based on a document analysis is fully supported by Powell and McCauley's (2002) analysis of discussions which took place in a number of national seminars on examining the doctorate. One additional factor discussed was the status and nature of the thesis and the viva: whether for example a satisfactory thesis but an unsatisfactory viva should lead to non-acceptance by examiners. Views of participants in the seminars were varied. The article distils from the discussions some 'Key factors in research degree assessment in the UK' (p. 109) and the 'Constitution of and roles within examining panels' (p. 112) which seem to have some general validity, but the final conclusion is that there is a need for greater transparency.

The emergence of new kinds of doctorate has also led to questions about comparability and consistency. Webb and Melrose (2014) discuss the 'creative writing doctorate', and Sharmini et al. (2015) and Wilson (2002) discuss the PhD which includes published work. Wilson charts the emergence of the publication-based thesis in one British University and the questions it raised, concerning comparability and whether the publication-based thesis is an 'easier' route. Sharmini et al. (2015) surveyed 62 supervisors in one New Zealand university from a range of disciplines, with a focus on ease or difficulty of assessment of theses containing publications. They first establish a taxonomy of such theses since there is a variation within this general category. They then examine the assessment process as seen by their informants, who for the most part 'used similar criteria for assessing publication-based theses as they would for monograph theses'.

Some, however, 'used their own criteria, extending the usual set' (p. 97). The authors suggest that this means it is time to 'revisit the 'universal' set of doctoral examination criteria and provide some new criteria for publication-based theses' (p. 98), but they leave this as a very general and open statement. The use of quotation marks for 'universal' also indicates their uncertainty about consistency in assessment of monograph theses, and their position remains unsatisfactory.

Kyvik (2014) compares examining procedures internationally, in Norway, Sweden, the United States of America, and the United Kingdom. The analysis is based on the reactions of academics in the latter three countries to their experience of examining Norwegian PhD theses, and the author's description of processes in the four countries. The purpose is to establish 'advantages and disadvantages of four different systems'. The conclusion is the rather unsatisfactory statement that 'different national assessment systems have their own pros and cons in terms of securing quality standards for PhD dissertations, and balancing different degrees of rigour and formality with efficiency in time use' (p. 152). The author asks if it is now time, in the light of internationalisation, to standardise assessment practices. The question he raises is still pertinent and still unanswered.

Analysis of criteria and factors affecting assessment

Criterion-referenced assessment necessarily involves interpretation by users of the meanings of words, and often leads to the use of examples to illustrate meanings, or to the discussion of criteria among assessors. When criteria vary across institutions, as is the case for the PhD, the complexity of the process increases. When criteria are not specific, this too creates a need for a shared understanding among assessors. All of this is highlighted in processes of QA.

It was in the context of growing attention to QA in the PhD examining process that Denicolo (2003) analysed criteria and how they are used in UK universities. She points out that criteria are 'most commonly couched at a very general level in terms of the work making an original contribution to knowledge and having the potential to be published in some form' (p. 86). In a questionnaire survey, asking examiners how they use criteria, she establishes considerable variation:

> the degree of consensus about the criteria for assessment is low, the degree of alternative interpretations of institutional regulations and procedures is relatively wide and the level of insecurity about (examiners') knowledge of general standards is high.
>
> (p. 90)

The question of views about and interpretation of criteria is pursued with two groups of examiners by Mullins and Kiley in Australia who, first, analyse the processes reported by experienced examiners (Mullins & Kiley, 2002) and then turn to inexperienced examiners (Kiley & Mullins, 2004). The former are, understandably, more confident in their judgements and use institutional criteria

sparingly and Mullins and Kiley (2002) analyse the examiners' own criteria and summarise first what examiners consider to be a poor thesis, often referred to as 'sloppy':

> 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;
> not being able to explain at the end of the thesis what had actually been argued in the thesis.
>
> (pp. 378–9)

A good thesis, by contrast, has 'a well-structured argument', 'a sensible and do-able question', 'a literature review that tells a story', an ability to communicate well, and 'reflection' including ability to critically assess one's own work. Interestingly, when the examiners were asked what makes an outstanding thesis, words such as 'creativity', 'artistic endeavour', and 'a well-sculpted piece of work' were used, especially by natural scientists.

Coming from a different academic tradition, in France, Dardy et al. (2002) analyse the statements in reports on oral examinations ('rapports de soutenance de thèse') to establish what is understood as a 'true thesis' both in the sense of the written text, the object, and with respect to the thesis which is propounded, the content:

> R [name of examiner] identifies, finally, one of the major qualities of the work presented: it is a question of a **true thesis** in the sense that the candidate truly presents a series of hypotheses and a point of view. He succeeds in defending and arguing for his thesis by using a range of explanations and by presenting a perspective on different approaches (history, micro and macro-economy, sociology of behaviour, biographical analysis…)
>
> (Dardy et al., 2002, p. 101 – our translation emphasis in original)

There are here familiar criteria for people in the anglophone tradition – the emphasis on hypotheses and arguments – but also a reference to a range of disciplinary perspectives which may not be so familiar.

Holbrook et al. (2004) come to similar conclusions about low- and high-quality theses and the criteria used to distinguish them. They also note that reports on low-quality theses are 'three times the length on average' of reports on high-quality theses, not least because they include detailed comments on how to improve a thesis, rather than condemning inadequacies. For example, a good quality thesis has 'expert use of the literature in design of the study and discussion of the findings', whereas a poor quality study has 'inadequate coverage or focus of the literature in relation to the study' (p. 117). The same team analysed

the specific comments in examiners' reports on the literature review (Holbrook et al., 2007), on references to theory in the thesis (2015), and formative comments provided by assessors (Holbrook et al. 2014).

Inexperienced examiners, according to Kiley and Mullins (2004) – by which they mean those who have examined fewer than three theses – tend to use institutional criteria more, and pay attention to 'technical' details although they still had a sense of what makes a 'good' thesis in their subject. Kiley and Mullins conclude that there are similarities and differences between experienced and inexperienced examiners:

> Both inexperienced and experienced examiners are similar in that they:
>
> - go about the assessment process in the same way (or at least with the same range of diversity);
> - use much the same criteria (looking especially for consistency and coherence, investigation of an interesting problem and insistence on high-quality presentation);
> - are impressed or put off by the same things, especially early in the thesis.
>
> However, inexperienced examiners differ from their more experienced colleagues in that they:
>
> - pay more attention to institutional guidelines;
> - tend to be uncertain about benchmarking, especially in regard to marginal theses;
> - have less experience of being a supervisor (and display a tendency to hark back to their own postgraduate experience).
>
> (p. 133)

Without distinguishing between experienced and inexperienced examiners, Kemp and McGuigan (2009) analysed the records of 283 examinations in a New Zealand university. They found no correlation among the initial, independent recommendations of examiners after their reading of the thesis. Nonetheless, the oral examination led to a compensation and increased the level of agreement, which suggests that the discussion among examiners, and the creation of a shared understanding, is crucial. Interestingly, on the other hand, Holbrook et al. (2008) also examined the degree of consistency among examiners, by analysing 2121 reports, and found 'an extremely high level of consistency in the examiner recommendations' (p. 45). Interestingly, even where recommendations differed, there was substantial similarity in comments, and adjudication by a committee to resolve differences drew on the comments rather than the recommendations. It is important to remember that, in the Australian system, there is no oral examination.

Finally, in a study of the 'report on the oral examination' ('rapport de soutenance de these') in France, an appendix of several pages lists many positive and

negative phrases used in reports to summarise the characteristics of theses and sometimes of their authors, based on close analysis of 14 such reports (Dardy et al., 2002). The authors deem it necessary to give three lists according to the three disciplines they studied in which a thesis is located: information and communication sciences, sociology, and linguistics.

In sum, it is evident, from literature which is dominated by research in anglophone countries, that formal, institutionally determined criteria are usually less than helpful, that experienced examiners in particular have their own criteria, and that consistency as reported by examiners or as analysed in their reports is reassuringly strong.

The oral examination

For many readers, it will be self-evident that an oral examination of some kind, in public or in private, is the final element of the PhD experience for students, supervisors, and examiners. That this is not the case everywhere is a significant reminder that the oral examination needs to be analysed in some depth. Such analysis was carried out in Australia a few years ago where there is no oral examination and where its purpose is questioned, although Kumar et al. (2020, p. 1) have recently said that Australian universities are 'increasingly introducing the oral'. Lovat et al. (2015) compared Australian, British, and New Zealand systems, the latter two having a 'viva voce' tradition, i.e., where the examination is held in private. They found, first, no consensus in the literature about the purpose of the viva. They also found that the assessment of the thesis is not changed by the oral examination, even though as Jackson and Tinkler (2001) found in analysing British university policy documents, the examination is considered an essential part of the process and could in principle lead to the candidate failing even if the thesis were satisfactory. In practice, this is rarely the case.

The tension between policy and practice is just one of the problems of the viva. Lovat et al. (2015) conclude however that despite the problems, the examination does serve the purpose of giving 'closure' to the PhD experience since in the best cases the candidate has the opportunity to meet their examiners, to discuss the work they have done over three or more years, and to exchange in a collegial fashion with the other scholars. We note that this also applies to public oral examinations in other traditions. On the other hand, there are some cases – it is not clear from the literature what proportion – when the candidate experiences the 'rite of passage' (Jackson & Tinkler, 2001; Lovat et al., 2015) as a 'horror story' (Sikes, 2017) where the experience might leave the candidate with diminished self-esteem and even a decision not to pursue an academic career.

The need to demystify and reduce such 'horror stories' is important not only for candidates but also for examiners (Sikes, 2017) since the viva can be an unpleasant experience for both. Examiners can however gain experience of the viva over time whereas the candidate has usually only one viva. It is therefore very important for a candidate to prepare well for their viva but this is very difficult because of the lack of transparency about the viva. Some authors have

nonetheless attempted to offer advice on preparation based on analysis of data from other examiners or their own experience (Sikes, 2017; Tinkler & Jackson, 2002; Watts, 2012). That the viva is an affective event and not a 'neutral, objective technology' (Sikes, 2017, p. 238) is well documented by Crossouard (2011) who concludes that 'there is reason to be concerned about the use of the *viva voce* for doctoral assessment, particularly for its unpredictability, its non-transparency (or perhaps the impossibility of transparency), and its unequal (and gendered) gaze' (p. 325) and 'the need for more reflexive questioning in academic circles about the subjectivities that are privileged within this assessment practice' (p. 328).

The fact that Crossouard suggests that transparency in the viva is perhaps not possible reminds us of its crucial characteristic of taking place in private. In some cases, there are no witnesses, there being present only candidate and examiners, although in recent years it has become common to appoint a convener or chairperson (Kumar et al., 2020). The viva can then lead to 'uncivil practices' (Bassnett, 2014) and become a 'horror story' (Sikes, 2017) for both candidate and examiners. Bassnett suggests that a change to a public event from the private viva is necessary but this would not necessarily resolve problems since the purposes of private vivas and public oral examinations are different. In an analysis of the public examination in France, up to around the year 2000, Dardy et al. (2002) emphasise its ritual nature, determined by a procedure which is 'rather precise and invariable' (p. 40). It is difficult to go further however since there is little literature with analysis of the public oral examination, an issue to which we return in the course of this book.

One issue which is likely to be common to both private and public examinations, but which is little discussed in the literature, is the question of the language used. As with other topics, the analysis focuses on anglophone universities and Carter (2012) investigates the experience of PhD candidates for whom English is an 'additional language'. This appears to be the only investigation and it too is intended to serve as a basis for preparation of candidates. She begins by saying the viva is 'slightly' more difficult for such candidates and ends with 'the surprising fact that all of these participants endorsed the oral examination suggests that they appreciate it as a final indisputable legitimisation of their doctorate from an English-speaking university' (p. 284). As in other investigations, the emphasis is on the ritual closure and, more critically, on success as 'experts in their subject in the English language' (p. 281). This latter point also appears in a tendency in non-anglophone universities to encourage the use of English in scientific discourse in general, and in oral examinations, even in public, in particular.

In summary, it is evident that, in terms of assessment per se, the (private) viva has little or no effect on the final decision about a thesis, but nonetheless has some positive characteristics of closure and ritual for most candidates. Yet the lack of transparency means there can be no meaningful preparation and, worse, some negative and even destructive experiences. The fact that this section is almost entirely devoted to the private viva is an indication of a substantial gap in the research literature which we hope in this book to begin to fill.

Examiners and their practices

In assessment systems which include an oral examination, PhD students meet their examiners, in public or in private, and there is necessarily a human relationship involved which as noted earlier can sometimes be traumatic for students and for examiners. As Sikes says:

> Doctoral examining is relational, emotional and ethical work and as such, despite academic cultural and institutional expectations and criteria as to what constitutes "doctorateness" (…), outcomes are influenced by subjectivities.
>
> (2017, p. 238)

This means that the selection of examiners is an important element of assessment as well as their practices once selected. Kiley (2009) analysed interviews with supervisors who, in anglophone countries, are influential in the selection process. The characteristics of an ideal examiner distilled from the data are someone who is experienced in supervising and examining, has expertise in the topic and methodology, an understanding of the university system in question and of the nature of the PhD, and is available when needed. Furthermore, they should have appropriate personality characteristics: 'high standards along with fairness, intellectual courtesy and generosity, and reliability' (p. 9). Like Sikes, Kiley also points out that there are important affective dimensions to examining.

Personality characteristics are no doubt inflexible and, for supervisor and candidate, not entirely knowable in advance. Other aspects of examining can however be learnt and, in principle, taught. Unfortunately, there appears to be very little if any training for examiners reported in the literature. Instead, there is analysis of how examiners act, and this could be the basis for training. Golding et al. (2014) identify eleven examiner practices for the benefit of students, 'what thesis students should know', which could equally serve novice examiners. Kyvik and Thune (2015) address the question of quality levels in PhD theses but also what effect on judgements of quality might be found in the characteristics of examiners. Mullins and Kiley (2002) use interviews with experienced examiners to analyse how they read a thesis, including the importance of 'first impressions' of the thesis, one of the 11 practices noted also by Golding et al. (2014). Kumar and Sanderson (2020) even analyse the effect of students' acknowledgements at the beginning of the thesis on examiners' judgements, while and Dally et al. (2022) analyse examiner feedback in the corrections required. Much of this material could be used as a basis for examiner training.

Less positive but equally important is the literature on the use and abuse of power in PhD assessment. Sikes (2017), cited earlier, give examples of power and abuse, and Lovat et al. (2004) draw on Habermas's thesis of 'ways of knowing' to analyse power discourse in examiners reports. In the editorial introduction to the collection which includes Lovat et al. (2004), Morley (2004, p. 91) gives a

warning which should be heeded, using a quotation which indicates that this is a long-standing problem:

> Doctoral assessment, as a social practice, is fraught with risks and uncertainties. Assessment in general is a major relay of power. The interpersonal aspects allow for distortions, disruptions and misrecognition. As a supervisor, examiner, and researcher, I am regularly reminded of what a precarious business it is and the complexity of the power relations involved in disciplinary expertise and academic judgements. (…) Anderson suggests the potential for abuse which exists today in Ph.D. examining is 'awesome'.

Benchmarking and value – and 'doctorateness'

In their analysis of the practices and perspectives of inexperienced examiners, Kylie and Mullins (2004) report that they regret the absence of a system of benchmarking to help examiners decide on the quality of a thesis. Benchmarks or standards have been linked to the question of QA but can also be seen as part of the more general discussion on defining 'doctorateness'. The association with QA began, in Britain, in the early 2000s (Shaw & Green, 2002; Morley et al., 2003). The British Quality Assurance Agency (qaa.ac.uk) published in 2020 a 'Characteristics Statement' for the doctoral degree which describes processes and procedures but does not define 'doctorateness' or refer to standards or benchmarks (Denicolo & Pak, 2013). The 'nonpartisan and objective research organization', NORC at the University of Chicago, has a 'doctoral benchmarking alliance' which allows universities to compare data on for example graduation statistics, but does not define a quality level or standard for a thesis/dissertation.

As for defining 'doctorateness', Wellington (2013) attempts this by analysing the possible purposes of the PhD, the regulations of institutions and opinions of examiners and supervisors and comes to the Wittgensteinian conclusion that

> instead of looking for the essential meaning of doctorateness we should look at various aspects, facets, behaviours, opinions, usages, enactments and activities that surround the doctoral endeavour, and use these to examine the evolving nature of doctoral study (…) to discern certain 'family resemblances'.
>
> (p. 1501)

This might also be an approach to the complexities introduced by variants of the traditional thesis-based PhD, such as professional doctorates, doctorates by publication or by portfolio. Anderson and Gold (2019) analyse the 'value' of the PhD asking supervisors in management education for their views, and compare these with an analysis of theses. They find similar characterisations to those identified in other studies but see a difference in the two sources of data. Thesis authors stress the applications and relevance of their work whilst supervisors emphasise originality, criticality, and academic worth. This article also argues for

a debate which includes consideration of 'research impact', an indication of how the purposes and nature of the PhD are changing, making the point that the Humboldtian tradition established the doctorate as an apprenticeship for an elite who would become university teachers whereas today, with many more people taking doctorate degrees, their work is important outside higher education, for society as a whole.

Ultimately, it may be inappropriate to try to define 'doctorateness' and standards or benchmarks, a position taken by Kelly (2017) who uses the concept of 'social imaginary' to 'argue for, and outline dimensions of, an imagined or culturally constructed PhD', and to do so by analysing how conceptualisations emerge not just within the academy but also in the wider cultural context, as revealed in 'stories, images and legends of society at large' (p. 5). She concludes, inter alia, that the Enlightenment idea of knowledge is still strongly present in the social imaginary, but that the growing presence of diverse and contradictory ideas is encouraging. Like many others, however, her analysis is of 'western' ways of thinking – and her reference base is in anglophone literature – and there is an obvious need to broaden the horizons to include analysis of other countries than 'western' and other literature than anglophone, which is one of the aims of this present book.

Research methodology

Since the focus of our study is on case-studies of specific universities and the experience of examiners and how they reflect on and conceptualise that experience, we chose semi-structured interviews as our main data collection instruments (e.g., Kvale, 2018; Spradley, 2016). In order to locate the interviews in each case study context, we also analysed the documents regulating the assessment process of each university. Furthermore, since we deliberately chose some cases where doctoral studies are well established and others where the PhD has been awarded only in relatively recent times, we analysed the relevant statistics.

The selection of case studies was 'purposeful' (e.g., Palinkas et al., 2015) in view of our decision to include both well-established and recent cases. Our research was helped both by our collaboration with universities in our previous book (Byram & Stoicheva, 2020) and by the involvement of Sofia University in international university networks. As editors and project leaders, we then selected cases from our knowledge of different countries, continents, and traditions which, without making any claim to 'representativity', would ensure a variety of experiences and a foundation for rich and stimulating case-study analysis together with the discussion of transversal issues, including the international standards and benchmarks.

A project with 13 case-studies from different countries and continents has to be flexible. There are more than 30 researchers involved, with different degrees and kinds of research experience. To ensure coherence and cohesion, we had online meetings, which had become the norm during the COVID-19 pandemic, although due to time-zone differences it was necessary to have two meetings for

each occasion, one for colleagues in Asia and Australia (and Europe) and one for those in the Americas (and Europe). In these meetings, the collection of documentary and interview data was discussed and a structure for the reporting of each case was suggested.

An interview schedule was prepared by the project leaders, based on their reading of the literature, on the explicit purposes of the project, and on their not inconsiderable personal experience of examining. This interview plan was piloted in December 2020 at the universities of Aveiro (Portugal) and Connecticut (USA). The schedule and piloting process were then discussed in an online meeting in January 2021 and further notes and thoughts were shared in an online document after the meeting.

As colleagues began to collect data, they had to change the schedule to suit their context. For example, the absence of a final oral examination in Australian universities meant that some topics for discussion in the schedule were not appropriate. Colleagues discussed changes with the project leaders and adapted the schedule to their context. Notes about changes in specific cases were shared with everyone in the project in an online document.

In each case, we tried to interview at least six people with examining experience. The main headings in the schedule were:

- Experience of the processes of examining
- Criteria used in assessment, and whether criteria differ for different disciplines
- The final oral examination (where applicable – see above): its format and purposes
- Standards: whether interviewees have experienced disagreements with other examiners and whether standards vary from university to university or country to country; whether grades are given and if so, how they are decided
- The presence of non-national languages in the dissertation/thesis and questions of how candidates who write in a second or third language are considered
- Experience of examining interdisciplinary projects
- The relationship between being a supervisor and an examiner, for example with respect to being a member of an examining group, committee or 'jury'
- Views on how the PhD has changed over time: for example, whether the students and their purposes have changed, and whether there have been specific innovations.

In the context of the pandemic, all interviews were conducted online. In non-anglophone universities, some interviews were conducted in English so that they might be accessible to other members of the research project, but only if interviewees were comfortable with this. Other interviews were conducted in interviewees' first language.

Transcription was at an 'ordinary language' level with the indications of nonverbal responses, where considered appropriate. Some transcripts were also

translated so that there could be a check of inter-coder reliability. Coding was carried out within the researcher groups of each case-study for the most part but occasionally with inter-coder reliability checks across cases, with groups sharing their transcripts with other groups. Coding was both deductive, based on the themes of the interview, and inductive from themes arising in the transcripts. The coding process was discussed and illustrated during one of the online meetings, and illustrative examples were shared in online documents.

Interviews lasted from 45 to 90 minutes. Interviewees were very generous with their time even though they were often senior people with many responsibilities, and in some interviews, there were indications that interviewees found the process interesting and were prepared to accord more time than they might have originally intended. Examining can be a stressful business for examiners as well as candidates. They are aware of the burden of responsibility in judging the dissertation/thesis and seldom if ever have training or even the opportunity to share experiences, and interviewees often seemed to enjoy the opportunity to discuss the issues. One interviewee articulated this, saying spontaneously towards the end of the interview – after she had delayed a scheduled meeting:

> It's a pleasure for me to talk about these things because I haven't given them much thought. I mean, you do it as they, as it comes. It's good when you start reflecting on your experience, and see it from a slightly different perspective, seek out bias.
>
> You are sometimes. I enjoyed the Webex meeting, as I said, I did enjoy as well because that makes you think about things that otherwise you don't have the time even to, to start thinking about.
>
> (SUE5)

The writing of case studies usually followed the agreed chapter structure discussed in a meeting but, as explained above, there was flexibility to fit the needs of each case, and some chapters are longer than others, for example, because there are current changes of significant interest.

The final stage of writing was the production of 'transversal' chapters, as we called them. In most cases, these were written by small teams from within the larger pool of authors of case-studies[3]. The themes for these chapters were almost self-defining in that editors and authors knew what we had to deal with, from our own experience as examiners, from the themes present in the first book and from the presence of these issues in the interviews. Initially, we intended that these chapters should come after the case studies, reflecting the chronology of the writing, but we decided that they should come before, in order for readers to benefit from overviews which might then help them better to understand the cases.

The cases are grouped mainly geographically, but we decided to use the term 'the West' metaphorically to put European, US American, and Australian universities together. This will no doubt seem controversial to some readers, implying a 'West versus the Rest' perspective. The rationale is that we are fairly

sure that there has been and is much interaction among PhD traditions in this group and yet there are differences which may go unnoticed. Putting them into one group is a comparative strategy in itself. At the same time, we are much less sure about traditions in other continents and decided to use a geographical grouping. Even this of course implies comparisons within continents and we could have simply put the chapters in alphabetical order by name of the university, for this is a study of universities and not countries, even less continents. Yet we wanted to give prominence to 'the Rest' and an alphabetical order would have diffused this. Whatever grouping we decided on would be debatable and doubtless criticised. Readers are in the final analysis able to read in any order they wish.

We invite readers, therefore, to use this book in different ways according to their needs and interests. They might wish to look at a transversal chapter of particular interest and then follow up on the references within it to specific cases. They might wish to look at a particular case study if for example they have been invited to examine in the country where the case is based. They might wish to look at cases in a particular region of the world. Whatever their approach, we trust they will find the accounts of examiners' experience illuminating.

Notes

1 We are very grateful to Stan Taylor for his help in establishing a corpus of work to review and analyse.
2 We meet here the first instance of a problem we find throughout the book: the translation of terminology. The term in the original language (e.g., защита (Bulgarian), 预答辩 (Chinese)) will inevitably have different connotations than 'defence', we cannot follow these in depth here. They will become evident in the details of each case-study. We opt for the generic 'defence' which can also be used for the 'viva' but have made the distinction here for purposes of emphasising the distance between the Anglo-Saxon tradition and others and its effect on research.
3 The exception is the chapter on standards on criteria, which was to be written by Nilza Costa and Michael Byram but, due to health issues, it had to be written only by Byram.

References

Anderson, V., & Gold, J. (2019). The value of the research doctorate: A conceptual examination. *International Journal of Management Education*, *17*(3), 1–14.

Bassnett, S. (2014) Cavalier attitudes lead to uncivil practices in vivas, *Times Higher Education*, available at: www.timeshighereducation.co.uk/comment/opinion/cavalier-attitudes-lead-touncivil-practices-in-viva/2013263.article.

Byram, M., & Stoicheva, M. (2020). *The doctorate as experience in Europe and beyond*. Routledge.

Carter, B., & Whittaker, K. (2009). Examining the British PhD viva: Opening doors or scarring for life? *Contemporary Nurse*, *32*(1–2), 169–178.

Carter, S. (2008). Examining the doctoral thesis: A discussion. *Innovations in Education and Teaching International*, *45*(4), 365–374. doi: 10.1080/14703290802377208.

Carter, S. (2012). English as an additional language (EAL) *vivavoce*: The EAL doctoral oral examination experience. *Assessment & Evaluation in Higher Education*, *37*(3), 273–284. doi: 10.1080/02602938.2010.528555.

Crossouard, B. (2011). The doctoral *vivavoce* as a cultural practice: The gendered production of academic subjects. *Gender and Education*, *23*(3), 313–329. doi: 10.1080/09540253.2010.508453.

Dally, K., Holbrook, A., Graham, A., & Lawry, M. (2004). The processes and parameters of fine art PhD examination. *International Journal of Educational Research*, *41*(2), 136–162.

Dally, K., Holbrook, A., Lovat, T., & Fairbairn, H. (2022). Supervisor perspectives on the 'end-stage' of the doctoral examination process. *Higher Education Research & Development 41*(2), 315–329. doi: 10.1080/07294360.2020.1847049.

Dardy, C., Ducard, D., & Maingueneau, D. (2002). *Un genre universitaire: Le rapport de soutenance de thèse*. Presses Universitaires du Spetentrion.

Denicolo, P. (2003). Assessing the PhD: a constructive view of criteria. *Quality Assurance in Education*, *11*(2) 84–91. doi: 10.1108/09684880310471506.

Denicolo, P. M., & Park, C. (2013) Doctorateness – An elusive concept? In: Kompf, M. and Denicolo, P.M. (eds) *Critical issues in higher education. Critical issues in the future of learning and teaching*, vol 8. Sense Publishers.

Golding, C., Sharmini, S., & Lazarovitch, A. (2014). What examiners do: What thesis students should know. *Assessment & Evaluation in Higher Education*, *39*(5), 563–576. doi: 10.1080/02602938.2013.859230.

Hoddell, S., Street, D., & Wildblood, H. (2002). Doctorates – converging or diverging patterns of provision. *Quality Assurance in Education*, *10*(2) 61–70.

Hodgson, D. (2020). Helping students understand PhD examination expectations: a framework and a tool for supervision. *Active Learning in Higher Education*, *21*(1), 51–63. doi: 10.1177/1469787417742020

Holbrook, A., Bourke, S., & Fairbairn, H. (2015). Examiner reference to theory in PhD theses. *Innovations in Education and Teaching International*, *52*(1), 75–85. doi: 10.1080/14703297.2014.981842.

Holbrook, A., Bourke, S., Fairbairn, H., & Lovat, T. (2007). Examiner comment on the literature review in PhD theses. *Studies in Higher Education*, *32*(3), 337–356. doi: 10.1080/03075070701346899.

Holbrook, A., Bourke, S., Fairbairn, H. & Lovat, T. (2014). The focus and substance of formative comment provided by PhD examiners. *Studies in Higher Education*, *39*(6) 983–1000. doi: 10.1080/03075079.2012.750289

Holbrook, A., Bourke, S., Lovat, T., & Dally, K. (2004). Investigating PhD thesis examination reports. *International Journal of Educational Research*, *41*(2), 98–120.

Holbrook, A., Bourke, S., Lovat, T., & Fairbairn, H. (2008). Consistency and inconsistency in PhD thesis examination. *Australian Journal of Education*, *52*(1), 36–48. doi: https://doi.org/10.1177/000494410805200103.

Irvine, G. (2012). On tenterhooks: A study of the time gap between submission and results of postgraduate theses. *Journal of Higher Education Policy and Management*, *34*(2), 185–186. doi: 10.1080/1360080X.2012.662742.

Jackson, C., & Tinkler, P. (2001). Back to basics: A consideration of the purposes of the PhD viva. *Assessment & Evaluation in Higher Education*, *26*(4), 355–366. doi: 10.1080/02602930120063501.

Johnston, S. (1997). Examining the examiners: An analysis of examiners' reports on doctoral theses. *Studies in Higher Education*, *22*(3), 333–347. doi: 10.1080/03075079712331380936.

Kelly, F. (2017). *The idea of the PhD. The doctorate in the twenty-first-century imagination*. Routledge.

Kemp, S., & McGuigan, C. (2009). Do PhD examiners agree on their recommendations? *New Zealand Journal of Educational Studies*, *44*(1), 47–53.

Kiley, M. (2009). You don't want a smart Alec': Selecting examiners to assess doctoral dissertations. *Studies in Higher Education*, *34*(8), 889–903. doi: 10.1080/03075070802713112.

Kiley, M., & Mullins, G. (2004). Examining the examiners: How inexperienced examiners approach the assessment of research theses. *International Journal of Educational Research*, *41*(2), 121–135.

Kumar, V., & Sanderson, L. J. (2020). The effects of acknowledgements in doctoral theses on examiners. *Innovations in Education and Teaching International*, *57*(3), 285–295.

Kumar, V., Sanderson, L. J., & Kaur, L. (2021). Investigating the role of convenors in the PhD viva. *Assessment & Evaluation in Higher Education*, 46(7) 1080–1091. doi: 10.1080/02602938.2020.1841092.

Kumar, V., Taylor, S., & Sharmini, S. (Eds.) (2023). *Doctoral examination: Exploring practice across the globe*. Routledge.

Kvale, S. (2018). *Doing interviews*. Sage.

Kyvik, S. (2014). Assessment procedures of Norwegian PhD theses as viewed by examiners from the USA, the UK and Sweden. *Assessment & Evaluation in Higher Education*, *39*(2), 140–153. doi: 10.1080/02602938.2013.798395.

Kyvik, S., & Thune, T. (2015). Assessing the quality of PhD dissertations. *A Survey of External Committee Members*, *Assessment & Evaluation in Higher Education*, *40*(5), 768–782. doi: 10.1080/02602938.2014.956283.

Lovat, T., Holbrook, A., Bourke, S., Fairbairn, H., Kiley, M., Paltridge, B., & Starfield, S. (2015). Examining doctoral examination and the question of the viva. *Higher Education Review*, *47*(3), 5–23.

Lovat, T., Monfries, M., & Morrison, K. (2004). Ways of knowing and power discourse in doctoral examination. *International Journal of Educational Research*, *41*(2), 163–177.

Morley, L. (2004). Editorial: Interrogating doctoral assessment. *International Journal of Educational Research*, *41*(2), 91–97.

Morley, L., Leonard, D., & David, M. (2003). Quality and equity in British PhD assessment. *Quality Assurance in Education*, *11*(2) 64–72. doi: 10.1108/09684880310471489.

Mullins, G., & 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. doi: 10.1080/0307507022000011507.

Murray, R. (2003). Students' questions and their implications for the viva. *Quality Assurance in Education*, *11*(2) 109–113. doi: 10.1108/09684880310471533.

Palinkas, L. A., Horwitz, S. M., & Green, C. A. *et al.* (2015) Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health*, 42, 533–544. doi:: 10.1007/s10488-013-0528-y.

Powell, S., & McCauley, C. (2002). Research degree examining – common principles and divergent practices. *Quality Assurance in Education*, *12*(2), 104–115.

Sharmini, S., Spronken-Smith, R., Golding, C., & Harland, T. (2015). Assessing the doctoral thesis when it includes published work. *Assessment & Evaluation in Higher Education*, *40*(1), 89–102. doi: 10.1080/02602938.2014.888535.

Shaw, M. & Green, D.H. (2002). Benchmarking the PhD: a tentative beginning. *Quality Assurance in Education*, *10*(2) 116–124. doi: 10.1108/09684880210423609

Sikes, P. (2017). And then he threatened to kill himself: Nightmare viva stories as opportunities for learning. *Qualitative Research Journal*, *17*(4), 230–242.

Spradley, J. P. (2016). *The ethnographic interview*. Waveland Press.

Tinkler, P., & Jackson, C. (2000). Examining the doctorate: Institutional policy and the PhD examination process in Britain. *Studies in Higher Education*, *25*(2), 167–180. doi: 10.1080/713696136.

Tinkler, P., & Jackson, C. (2002). In the dark? Preparing for the PhD viva. *Quality Assurance in Education*, *10*(2), 86–97.

Trafford, V. (2003). Questions in doctoral vivas: views from the inside. *Quality Assurance in Education*, *11*(2) 114–122. doi: 10.1108/09684880310471542

Watts, J. H. (2012). Preparing doctoral candidates for the viva: Issues for students and supervisors. *Journal of Further and Higher Education*, *36*(3), 371–381. doi: 10.1080/0309877X.2011.632819.

Webb, J., & Melrose, A. (2014). Understanding the value and the impact of the 'Shock': Examining the creative writing doctorate. *New Writing*, *11*(1), 134–148. doi: 10.1080/14790726.2013.849744.

Wellington, J. (2013). Searching for 'doctorateness'. *Studies in Higher Education*, *38*(10), 1490–1503. doi: 10.1080/03075079.2011.634901.

Wilson, K. (2002). Quality assurance issues for a PhD by published work: a case study. *Quality Assurance in Education*, *10*(2), 71–78. doi: 10.1108/09684880210423555.