

‘It’s a PhD, not a Nobel Prize’: how experienced examiners assess research theses

GERRY MULLINS

Adelaide Graduate Centre, Adelaide University, Australia

MARGARET KILEY

Centre for the Enhancement of Learning, Teaching and Scholarship, University of Canberra, Australia

ABSTRACT *Research to date on the examination process for postgraduate research theses has focused largely on the deconstruction of examiners’ reports. This article reports on a study of the processes that experienced examiners go through, and the judgements they make before writing their reports. A sample of 30 experienced examiners (defined as having examined the equivalent of at least five research theses over the last five years), from a range of disciplines in five universities was interviewed. Clear trends emerged with regard to: the criteria used by examiners and the levels of student performance expected by them; critical judgement points in the examination process; the examiners’ perceptions of their own role in the process; the influence on examiners of previously published work, the views of the other examiner(s) and their knowledge of the student’s supervisor and/or department, and the level of perceived responsibility between student and supervisor.*

Introduction

Despite the fact that universities have been assessing doctoral and master’s theses for many years, there has been little research done on the processes involved in that assessment (Tinkler & Jackson, 2000). This situation is in striking contrast to the situation with undergraduate assessment. Academic staff seeking assistance with the assessment of undergraduate students can be provided with a wealth of advice—all based on extensive quantitative and qualitative research, and well-documented case studies leading to clearly formulated guidelines. However, such a body of knowledge is not available to assist examiners in the assessment of PhD theses. In the absence of a clear, well-researched understanding of the examination process, anecdotes, generally of the traumatic kind, abound among the student body. Many supervisors are poorly placed to refute these anecdotes as often they are able only to speak from their own, sometimes narrow, experience.

In Australia, PhD awards are based on a written thesis reporting the results of a three to four year research programme. An oral defence of the thesis is only available at a few Australian universities, and is generally at the request of one of the examiners. The request usually results from a level of ambiguity within the thesis, or lack of certainty on behalf of the examiner as to the student’s grasp of a particular issue. Although there is a steady increase in the number of courses for research students, performance in these courses is not taken into

account when assessing the student for the award. Given the significance of the single piece of written work in assessing a student for the award of PhD (or Research Master's), there has been surprisingly little research on the way in which examiners make judgements about the quality and quantity of the research work and the way it is reported. Nor do we have a clear understanding of how examiners undertake the assessment process, what it is they believe themselves to be doing, and why they undertake the time-consuming and often stressful task of examining theses. While, in other countries, the significance of the written thesis may vary in relation to other components of the assessment (e.g. see Tinkler & Jackson (2000) on the viva in PhD assessment in the UK), questions about how examiners go about the task of assessing the thesis itself remain.

The research project on which this article is based set out to answer two questions:

- Is it possible to define one aspect of the pedagogy for postgraduate learning, i.e. assessment?
- What advice might there be for students, supervisors, examiners and institutions related to the examination of postgraduate research theses?

More specifically, the project addressed the following issues:

- What steps do examiners go through in the process of reading a thesis?
- What criteria do examiners use to assess a thesis? Are these criteria derived from institutional policies or are they based on the individual examiner's understanding of what is required for a PhD or a master's by research?
- Do examiners use different criteria for different groups of students (e.g. international students from non-English speaking backgrounds)?
- Are there influences on the examiners arising from their knowledge of the university, department or supervisor?
- What evidence do examiners collect as they read a thesis with a view to the formulation of their final written reports?
- Are there critical points in the process of making judgements about a thesis which significantly influence the examiner's final evaluation of the thesis?

Previous Research

Nightingale (1984) analysed examiners' comments on 58 theses submitted at an Australian university. She concluded that examiners needed more detailed criteria than the statement that a PhD thesis should make 'a substantial and original contribution to knowledge'. Nightingale recommended that:

- the criteria by which research degrees are assessed be clarified so that examiners receive an adequate description of the range of degrees awarded by the institution, a description of how they differ from each other, and of the specific requirements of the different programmes;
- examiners be aware of the examination processes of the particular institution and of the options open to the examiners.

Hansford & Maxwell (1993) examined 255 examiners' reports relating to 125 part-time, external master's degrees in Education. They found that examiners focused on format and presentation, the literature review, the theoretical/philosophical framework of the thesis and the problem of unsubstantiated and/or over-generalised conclusions. However, they seemed

to comment little on research questions and design, data collection and analysis, and on the supervision itself.

Two more recent Australian studies, Ballard (1996) and Johnston (1997), continued the analysis of examiner's reports. Ballard (1996) noted that the traditional qualities of a successful thesis—originality, scholarship and advancement of knowledge—are 'transformed by the examiners into the less lofty expectations of "imagination", "competence" and "mastery"' (p. 2). She concluded that:

Examiners assume PhD candidates are still apprentices in the profession of research in their discipline; and so their theses are judged in terms of current competence and future promise as academic colleagues. If there are particular problems, then the examiners regard the department, the supervisor and the candidate as all being potentially implicated; and if there are remarkable achievements, the recognition likewise extends beyond the performance of the individual candidate. Similarly, the examiners themselves are conscious that their own reputation is being judged through the quality of their reports. (pp. 13–14)

Johnston (1997) stressed the communications aspect of the examination process. She questioned the assumption that examiners are 'first and foremost experts in the field who will judge work accordingly' (p. 345), and suggested that examiners approach the reading of a thesis just like a reader of any new piece of writing:

Examiners require all of the normal forms of assistance which should be provided to any reader. They appreciate work which is logically presented, focused, succinct, summarised and in which signposts are used to help readers to understand the path they are taking through the work ... One of the problems with work that is poorly presented is that the examiner tends to lose confidence in the candidate and can become suspicious that there are deeper problems of inadequate and rushed conceptualisation. (p. 345)

Johnston, like Nightingale, called for a more transparent examination process and for more explicit and detailed consensus about the required standard of a PhD.

Pitkethly & Prosser (1995) used examiners' reports on 74 PhD theses in an attempt to identify the extent to which examiners' comments placed theses in an international context. They found strong similarities and only minor differences between Australian and international examiners.

These studies are illuminating, but their focus on the examiners' reports restricts the insights they might have gained into the assessment process. By the time an examiner has written their report, they have gone through several stages of an extended process of evaluation and made a number of important judgements about the quality of the thesis—none of which will necessarily be included in the final report.

As part of a study funded by the Economic and Social Research Council (UK), Phillips (1992) reported on interviews with 58 academic staff about the criteria they applied as external examiners. Phillips reported considerable agreement concerning the format and general content of theses, and about what is expected with regard to standards. While technical proficiency, originality and conceptual development were important, for these social scientists, 'originality' was accepted as meaning 'creative' or 'significant'. Delamont *et al.* (2000), as part of a UK study of the socialisation of doctoral students, also interviewed 15 examiners, mostly from the areas of urban planning and developmental studies, about what they looked for in a PhD. They were struck by the 'indeterminacy' of the criteria listed by their respondents; for example, 'coherence', 'rigorous argument', 'meaty', 'thoroughness'.

TABLE I. Sample by gender and discipline

	Science	Maths/ Engineering	Social science	Humanities	Total
Male	11	1	4	2	18
Female	3	2	5	2	12
Total	14	3	9	4	30

They concluded, ‘While there are technicalities which must be correct, the real role of the examiner is to judge whether the student has mastered appropriate indeterminate skills and displayed the right indeterminate qualities’ (p. 41).

Finally, Winter *et al.* (2000) reported on the criteria used by UK examiners in the process of assessing ‘borderline’ theses. They were concerned with the possible discrepancies between the criteria for ‘academic’ and ‘practice-based’ or ‘professional’ doctorates. To establish a reference point, they sent a questionnaire to 31 examiners from a range of disciplines and institutions, asking them to reflect on the criteria they used to decide whether a PhD was acceptable, and on what they meant by ‘original’ and ‘publishable’. The collated criteria described by their respondents are extensive (Winter *et al.*, 2000, pp. 32–35). They concluded ‘that a PhD ought to:

- be a report of work which others would want to read;
- tell a compelling story articulately whilst pre-empting inevitable critiques;
- carry the reader into complex realms, and inform and educate him/her;
- be sufficiently speculative or original to command respectful peer attention (p. 36)’.

However, the intention of Winter *et al.*’s study was merely ‘to find out, in a sense, the scope of the problem’ (p. 31) rather than to map definitively their respondents’ criteria; nor did they have the opportunity to explore, beyond the questionnaire, their respondents’ views.

As a first step in determining what examiners are looking for in a research thesis, we have interviewed a sample of 30 experienced examiners. Experienced examiners, defined for the purposes of this study as having examined the equivalent of five theses over the past five years, were chosen for the initial stage of this research for three main reasons. The first was that they were in a position to speak from considerable experience based on the number of theses they had examined, compared with perhaps only one or two theses examined by inexperienced examiners. The second reason was that we considered that these examiners, by dint of their experience, might more easily articulate their position with regard to the process of judgement employed. The third reason was that these examiners might feel more confident about sharing their experiences with us and being prepared to suggest further aspects of the research into which we should delve.

Method

The sample for this study was drawn using gender, discipline and the methodological paradigm of the examiner as the criteria (see Table I). For this study, all the interviewees were staff members of a range of Australian universities, and we made clear to them that our interest was in the examination of the traditional PhD, as typically submitted to Australian and UK universities.

Face-to-face interviews were selected as the most appropriate method for the project. Pilot interviews indicated that it was necessary to press examiners to be explicit about what

they did in the process of examining a thesis and why they did what they did. In the interviews, we sought to identify their understanding of why examiners thought they were chosen and why they accepted the request. We asked how they went about examining the thesis and what they were thinking at each stage of the process. In addition, we asked the interviewees to identify any biases or influences of which they were aware and the expectations they held of the thesis, the student and themselves. The interviews tended to begin with broad generalisations about the process, but, with encouragement, examiners made specific statements, and gave examples and explanations of what they were looking for in a thesis, and what counted as evidence of attainment of their criteria. Our experience with the interviews of experienced examiners indicates that this approach yields useful information relevant to the project goals that would not be available through questionnaires or other forms of questioning.

As researchers, we set out to gain an understanding of what led respondents to develop their conceptions of an acceptable research thesis, and the context within which they developed their conceptions (Johansson *et al.*, 1985). The decision to adopt this research approach had two interesting implications. Several interviewees commented that, while they thought they were answering the questions posed ‘honestly’, they recognised that they almost certainly held biases of which they were not specifically aware. For example, ‘I know I have more biases than I would like to think I have, but I try to be open’ (SocSc/Female/3). In addition, several interviewees asked us how we knew that what they were saying was true. We explained that we were trying to gain an understanding of what they thought they were doing during the examination process and why. In other words, we were taking at face value what they told us.

The research can be identified as having four stages. In Stage 1 the questions were developed using the literature and discussions with two very experienced examiners (one science and one social science). Stage 2 involved interviewing 12 experienced examiners at Adelaide University. The method of semi-structured interviews was adopted (Brenner, 1985; Fontana & Frey, 1994), whereby the questions were asked in an order and in ways that fitted with the manner in which the interview was progressing. Each interview was taped, with the agreement of the interviewee, and then summarised, including verbatim quotations, from the tape. The summaries were sent to the interviewees for confirmation.

A key concept in this form of interpretative research and analysis is the notion of ongoing analysis (Denzin, 1994). Stage 3 involved analysing the data and revising some of the questions. ‘Analysis during data collection lets the field worker cycle back and forth between thinking about the existing data and generating strategies for collecting new—often better quality—data’ (Miles & Huberman, 1984, p. 49). This process allowed for the data to be ‘both the evidence and the clues’ (Bogdan & Taylor, 1975, p. 73). Stage 4 comprised interviews with staff from other universities. This data, using a similar method to that described above, was analysed in light of the first round of interviews and analysis.

We analysed the data in four quite different ways. First, we analysed the responses to the same question, looking for similarities and differences in the responses; for example, the responses to the question, ‘Are you influenced by seeing that the student has published during candidature?’ The second form of analysis was to look for metaphors in the responses as an indication of what respondents meant by a ‘good’ or ‘poor’ thesis. It is argued that interviewees used metaphor to describe what they found to be difficult concepts. The third form of analysis was to search for the words and phrases that examiners used to describe what they were looking for in a good thesis and to cluster them into categories. The fourth form of analysis was to check for consistency within each response. We were interested to see

whether what examiners indicated was the purpose of a PhD was consistent with the evidence they used to decide on the quality of a thesis.

Results

Why Examiners Thought They Were Chosen

We were interested to know why experienced examiners thought they were chosen to examine, as we hypothesised that this would be an indicator to their frame of mind when examining. The criteria used by Australian universities to select examiners vary widely: some universities have formal requirements that examiners have both a PhD and research and supervisory expertise in the area; others have few formal criteria beyond expertise in the discipline.

Most interviewees believed that they were asked to examine because they were experts in their field and so they had a good sense of the standards for the discipline. This belief manifested in several ways; for example: ‘Other supervisors want their students to be “known” by an expert in the field’ (Maths–Eng/Female/20), or: ‘I consider that when I select an examiner it is a stepping stone for the student’ (Maths–Eng/Male/21). The expertise of examiners was also demonstrated by the student who frequently cited the examiner’s research in his/her PhD: ‘[My] work is often cited by students and so, when looking for an examiner, as one does when looking for a referee for a journal, one will look at the references. In a fairly small subdiscipline, the supervisor would probably know the work in that area’ (Sc/Male/4). As we will demonstrate later, the result of the examiner’s work being known to the student and the supervisor is one of the biases that we suggest is demonstrated in the way that experienced examiners responded to the question regarding methodological stance.

Some interviewees believed they were qualified in areas where few others were qualified to examine. For example, in one social science discipline the examiner suggested that there were very few women with a PhD and so she was asked to examine a wide range of feminist work within the discipline. Another experienced examiner in science suggested that most graduates in her subdiscipline moved immediately into industry and so there were few academics in this area who could examine: ‘In my area in Australia there is not a lot of choice!’ (Sc/Female/12).

Finally, some interviewees thought they were asked to examine because they were considered to be sympathetic toward the student’s situation: ‘I think people will expect me to be understanding of the student’ (SocSc/Female/27).

There was virtually unanimous agreement that inexperienced examiners needed to be avoided at all costs with theses that might pose concern. Two reasons were suggested for this avoidance. One was that inexperienced examiners had little understanding of the constraints upon PhD students and the understanding required of their situation: ‘I suspect that one of the reasons inexperienced examiners are so tough is that they have not experienced their own students being scrutinised’ (SocSc/Female/27). The other reason was that the inexperienced examiner had little in the way of comparison of standards of work: ‘Inexperienced examiners are the most dangerous examiners—mainly because they often have a sample size of one on which to judge—their own—and so anything that comes to them will be judged on whether it is better or worse than their own’ (Sc/Female/14).

Why Examine?

Given the time and effort examiners put into examining in return for a very low honorarium, why is it that they accept the request to examine in the first place? Two main reasons were

given: more than half of all the interviewees suggested that they examined out of a sense of duty, but this duty was defined in three different ways. First, there was a sense of duty regarding maintaining standards within the discipline: 'You are asked to maintain the standards because of your own professional expertise' (Sc/Male/13). This was the most common response, followed by the belief that one of the roles of an academic is to examine theses: 'It plays a role in the education process and if examiners didn't do it, the whole system wouldn't work' (Sc/Male/5). The third way that examiners discussed duty was their duty to their students, or the 'quid pro quo' concept: 'It's ... a reciprocal obligation from having one's own students examined' (Sc/Male/26). There was often a sense of 'the more students one has, the more one has to examine'. As one examiner said, 'I have eight students at the moment which means I need 16 examiners soon ... so I need to reciprocate' (Sc/Male/19).

However, some examiners also cited other reasons for examining. These reasons included the excitement and interest involved: 'The enjoyment of a really good thesis showing lots of promise' (SocSc/Male/15), as well as access to state of the art research. Moreover, a thesis usually includes a level of detail not included in examiners' day-to-day professional reading: 'A good thesis includes a lot more detail than in articles' (Maths-Eng/Female/18).

Influence of the Examiner's Methodological Stance

As researchers, we were particularly interested in determining the strength and influence of the examiner's methodological paradigm on the process of examination. We had hypothesised that paradigm would be one of the major differentiating factors in the sample, and one of the main influences on the examination process. While the findings indicate that certainly paradigm was a significant disciplinary difference, we were somewhat surprised by the findings related to the influence of paradigm on examination.

In all but two cases, earth sciences and environmental science, examiners working in the sciences and mathematics/engineering commented that they looked for 'good science' when examining. Good science, according to the responses, can be summarised as 'A pertinent literature review, clear hypothesis, do-able problem, sound data analysis and methodology, and justifiable conclusions'. In the humanities and social sciences areas, despite actively seeking out interviewees in subdisciplines where one might expect a strong paradigmatic influence on examination, other than in two cases, both in education, we found that most experienced examiners described themselves as 'eclectic', 'catholic', or 'generalist'. Several suggested that they had become far less 'doctrinaire' with experience, and considered themselves more able to examine across a range of paradigms because of their experience. However, what most examiners did assert strongly was that they looked to see that students were consistent and that they had actually done what they said they were going to do rather than adhere to a particular paradigm or methodology: 'I try in my reading of theses to understand where the student is coming from. Even if I don't agree with the perspective they have, or if there are gaps, I try to see it from their eyes and whether they have been true to what they set out to do' (SocSc/Female/27).

When further questioned, these experienced examiners admitted that generally their work was well known within either a particular subdiscipline or methodological approach, and so it was unlikely that they would be asked to examine in a paradigm with which they were not sympathetic.

Time Devoted to Examining

The interviewees in this study were, by definition, senior academic staff—most were full professors. Often they were the heads of their departments/schools; leaders of large research

teams; the editors of a journal, or on the review panels of several journals; members of national research grants committees and on several institutional committees. In spite of heavy workloads, most indicated that they spent the equivalent of three or four days fulltime examining a thesis, often over a period of two to three weeks. There were many accounts of people carefully planning to set aside extended periods of uninterrupted time to do this task, often at night or over weekends—none reported scrutinising theses at their workplace. The process varied—some examiners read the thesis from cover to cover three or four times, others only once—but in all cases they clearly indicated that the task was demanding but thorough. As one interviewee indicated, ‘You’ve got a lot of somebody’s work. On the one hand it is crucial to do justice to that work, but it is also important to ensure that it should mean a lot to get a thesis’ (SocSc/Male/29). The time and effort put into the examination process is a credit to examiners, and a comfort to postgraduate students.

Experienced Examiners Expect the Thesis to Pass

Of even more comfort to postgraduate students is the reluctance of examiners to fail a thesis. From our 30 experienced examiners (who had examined more than 300 theses over the last 10–15 years), there were only 10 reports of a failed thesis. There are several reasons given for this reluctance. Primarily, it is the examiners’ realisation that the thesis represents three to four years of effort by a talented student, and that its production has been an expensive process in terms of resources and other people’s time: ‘If the student is any good and the supervisor any good then you shouldn’t fail a PhD. There should be enough “nous” around to guide the student in a way that he/she wouldn’t fail’ (Sc/Male/10).

Another reason examiners will do everything they possibly can to avoid failing a thesis, or asking for a substantial rewrite, is that they realise that this will require a substantial amount of work for the examiner, the student and often the supervisor: ‘A poor thesis causes me sleepless nights as I know how much work and effort is involved’ (Hum/Female/6).

How Examiners Work through the Thesis

We were interested to know whether experienced examiners read the thesis from cover to cover, as one might read a book, or whether they treated a thesis differently. As one might expect, different examiners approach the task differently, but most examiners begin by reading the abstract, introduction and conclusion to gauge the scope of the work, and by looking at the references to see what sources have been used and whether they need to follow up on any of them. They then read from cover to cover, taking detailed notes, and finally go back over the thesis to check on whether their questions have been answered or whether their criticisms are justified. However, four examples give a good indication of the range of ‘reading styles’:

- A (Hum/Male/17) sets aside time to read the thesis. He checks who is in the references to see that the writers are there who should be there. Then he reads slowly, from the beginning like a book, but taking copious notes.
- B (Sc/Male/22) reads the thesis from cover to cover first without doing anything else. For the first read he is just trying to gain a general impression of what the thesis is about and whether it is a good thesis—that is, are the results worthwhile. He can also tell how much work has actually been done. After the first read he then ‘sits on it’ for a while. During the second reading he starts making notes and reading more critically. If it is an area with which he is not very familiar, he might read some of the references. He marks typographical

errors, mistakes in calculations, etc., and makes a list of them. He also checks several of the references just to be sure they have been used appropriately.

- C (SocSc/Female/27) reads the abstract first and then the introduction and the conclusion, as well as the table of contents to see how the thesis is structured; and she familiarises herself with appendices so that she knows where everything is. Then she starts reading through; generally the literature review, and methodology, in the first weekend, and the findings, analysis and conclusions in the second weekend. The intervening week allows time for ideas to mull over in her mind. On the third weekend she writes the report.
- D (SocSc/Male/15) reads the thesis from cover to cover without marking it. He then schedules time to mark it, in about three sittings, again working from beginning to end. At this stage he ‘takes it apart’. Then he reads the whole thesis again.

The questions that examiners have in mind as they read include the following:

- How would they have tackled the problem set out in the abstract and the title?
- What questions would they like answers to?
- Do the conclusions follow on from the introduction?
- How well does the candidate explain what he/she is doing?
- Is the bibliography up to date and substantial enough?
- Are the results worthwhile?
- How much work has actually been done?
- What is the intellectual depth and rigour of the thesis?
- Is this actually ‘research’—is there an argument?

First Impressions Count!

An overwhelming conclusion from this research was the extent to which examiners’ first impressions counted. These first impressions were not irreversible, but they did influence the examiner’s frame of mind for the rest of the thesis. Experienced examiners decide very early in the process whether assessment of a particular thesis is likely to be ‘hard work’ or ‘an enjoyable read’. However, several examiners commented on how they were careful not to be overly influenced as they read through the rest of the thesis, and some recalled instances when a thesis did not live up to initial expectations or when a candidate rescued what looked like a poor thesis.

The initial impressions of the quality of the thesis are usually formed by the end of the second or third chapter of the thesis—often by the end of the literature review. The authors are aware that the ‘traditional’ format of a thesis—introduction, literature review, methodology, results, conclusions—is no longer universally accepted as appropriate (Bruce, 1994). However, none of our interviewees indicated that the theses they read departed much from this model. Two examples are typical of the influence of first impressions:

- ‘A good indicator is the way the candidate reviews the literature and their overall grasp of what’s going on’ (Sc/Male/3). If it looks as if the student grasps the problem then this examiner reads the rest with much more of a sympathetic view and he feels he can relax. If chapter 2 is not good, then he reads the rest much more critically.
- ‘It is unusual that if someone does a poor job of the literature review that they will suddenly improve, or vice versa’ (Sc/Male/5). This examiner looks for originality, a good understanding of the subject, and at the quality of the literature review with interpretation.

Other examiners form this first impression in the course of their initial scoping of the assessment task. For example, one examiner (SocSc/Female/8) reads the table of contents, the first chapter and then the last chapter because she believes the last chapter should take off from where the first chapter ends. This reading simultaneously gives her a sense of whether the student has a 'thesis', and whether the work is of good quality. Another examiner (SocSc/Female/27) described her experience with a good thesis as follows: the first chapter allayed her fears, given that it was on a topic that she had initial reservations about; and the theoretical chapter (chapter 2) was masterly in that the student had covered all the aspects one would hope to see and had managed constraints, etc. The examiner had a sense of the design and conduct of the research, along with the appropriate modifications, by the end of chapter 2. In general, experienced examiners seek to be assured that 'This person obviously knows what they are on about' (Science/Female/12).

What Makes a Passable Thesis?

Any valid assessment of a learning process might be expected to show consistency between the goals of the learning process and the evidence used to judge whether those goals are achieved by the student. We were interested, therefore, to compare our interviewees' statements about the purpose(s) of the postgraduate experience with the criteria they used to assess the thesis. Typically, experienced examiners saw the purpose of a postgraduate research programme as being either (a) the production of a thesis with given characteristics, or (b) the development of the skills and attitudes necessary for the student to operate as an independent researcher. Often, both (a) and (b) sat comfortably side by side in the mind of the interviewee. While not all interviewees were asked to define the goals of a PhD programme, the 23 who did so also were quite consistent in seeking appropriate evidence to assess the achievement of those goals.

Many of the interviewees started by first stating what they thought made a poor PhD, and then defining a good or passable one. One of the most common descriptors of a poor thesis, across all disciplines, was 'sloppiness'. Sloppiness might be demonstrated by typographical errors, or mistakes in calculations, referencing and footnotes. The concern with sloppiness was that examiners considered it was an indicator that the research itself might not be rigorous and the results and conclusions could not be trusted. For example, 'If there are mistakes in calculations then that starts one thinking. If you find something suspicious in a thesis then you start to read it differently' (Sc/Male/22).

Several examiners also commented on how they can be easily irritated by inattention to detail:

I give my students strong advice on how not to 'flip' an examiner from 'reasonable' to 'unreasonable' by having irritating things in the thesis such as typos and other careless textual mistakes that indicate lack of attention to detail. Once flipped (and I am aware of this happening), I am irritated and I have to work very hard at overcoming this irritation and not letting it influence my view of the thesis, although this is not easy. (Sc/Female/14).

Characteristics of a poor thesis were:

- 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.

On the other hand, a term used frequently to describe positive theses was ‘scholarship’, described by interviewees from all disciplines as originality, coherence, and a sense of student autonomy or independence: ‘The student makes the ideas their own’ (Hum/Male/17); ‘The original use of a concept or theoretical framework’ (SocSc/Female/7); ‘To open up new areas’ (Sc/Male/13); ‘The student has done what they said they would do’ (SocSc/Male/29).

The development of a well-structured argument was highly valued in a thesis. Within this term we clustered argument, conceptualisation, conclusion, design, logic and structure. Comments included: ‘Logical progression of ideas, work and presentation’ (SocSc/Male/15), and ‘Higher level thinking and analysis’ (SocSc/Female/28). Examiners also sought ‘The selection of a “real” problem’ (SocSc/Male/11); ‘A sensible, do-able question’ (Sc/Female/12); along with ‘A literature review that tells a story’ (Sc/Male/5). The student’s ability to communicate was crucial: ‘The student takes you on a journey’ (SocSc/Female/27), and ‘Succinct writing without speculation’ (Maths–Eng/Male/21).

Most examiners looked for sufficient quantity as well as quality of work, with the frequent use of the word ‘substantial’. In the sciences, this was often defined in terms of the number of journal articles likely to arise from the work. This ranged from two to four good journal articles—not that the student had to have published these, but that there was sufficient material in the PhD to allow for these articles to be published. Linked with this concept was that of publishability; that is, the PhD was of a sufficient standard that parts of it, or the whole, could be published.

‘Reflection’ was another general characteristic that experienced examiners sought: ‘They make a critical assessment of their own work’ (Sc/Male/13), and ‘They are critical of their own argument’ (Hum/Female/24). Examiners also looked for students who were able to work their way through problems: ‘How they recognise and deal with contradictions’ (SocSc/Male/15), and ‘An appreciation of what the research means’ (SocSc/Female/28).

What Makes an Outstanding Thesis?

Interviewees were asked to comment on what they thought set apart a ‘good’ PhD from a standard or passable PhD. There was considerable unanimity across the disciplines with regard to these characteristics, and one of the unifying responses was the use of the artistic metaphor. For example, words and phrases such as the following were used, particularly by scientists, to describe a good PhD:

- ‘an artistic endeavour where the student is designing the work and there is elegance of design, of the synthesis, and executions’ (Sc/Male/22);
- creativity;
- design—where it all fits together;
- elegant;
- a well-sculpted piece of work.

The use of the artistic metaphor extended to such terms as ‘élan’, ‘passion’, ‘excitement’ and ‘sparkle’. Clearly, these experienced examiners believed there was a level of ‘art’ involved in producing a particularly good thesis.

Examiners were looking for students who exhibited a sense of confidence in the way they dealt with the material and a level of sophistication in the way they presented their argument. Also, as one examiner said; ‘The outstanding PhDs have beautifully conceived ideas that open up a new area or really answer an important question, and are critical of previous work in the area or make a critical assessment of their own work’ (Sc/Male/13). Or, put another way:

All PhDs are not equal and yet most get through. You form an impression that it is OK but not dazzling. This is often when the student applies standard theories in a rather pedestrian way. It’s not wrong, and you can’t fail it, but it’s not dazzling. In other cases you can see that the material is taken and used originally at every level—methodology, literature review, etc. Right from the beginning it makes you see an area that you thought you knew in a way that you hadn’t thought about before. (Hum/Male/8)

Influences on the Examiner

Interviewees were asked a number of questions about the extent to which they were influenced in their judgement by factors such as:

- statements about the criteria and/or specified levels of performance forwarded by the student’s institution;
- the views of the other examiner(s);
- work previously published by the candidate;
- their knowledge of the student’s supervisor, department and/or institution.

Answers regarding these factors varied. However, a striking characteristic of experienced examiners was their confidence in their own judgement regarding the quality of postgraduate research.

It is not uncommon for institutions to ask examiners to assess theses from that institution according to certain criteria which are much more specific than the traditional ‘original contribution to knowledge in the discipline’ (Tinkler & Jackson, 2000). Some examiners did indicate that they checked the guidelines in some way; for example, they ‘took them under advisement’ (Hum/Male/8). Others said that, before finalising the report, they did a ‘reality check by looking at the criteria’ (Sc/Male/4). Finally, some examiners reported that they paid particular attention to the guidelines when making their final summative judgement, especially in discriminating between recommending resubmission as distinct from recommending changes to the satisfaction of the supervisor(s) or the head of department (Science/Female/14). However, only a third of our examiners took institution-specific criteria into account in assessing the thesis. Most examiners wrote their reports in the form requested by the institution, but, when it came to the point of making a judgement, they regarded themselves as the arbiters of an acceptable thesis: ‘No first rate researcher is without a belief that they understand the standards in that field and can recognise excellence in that field ... So if you ask me to examine, you are going to get [my] standard’ (Science/Male/4).

Other examiners. A similar independence of judgement is evident in regard to experienced examiners’ attitudes to other examiners. In some cases, interviewees thought that they were prevented by confidentiality regulations from contacting the other examiner(s). Most, how-

ever, would never wish to approach the other examiner(s), although they did not mind being contacted to discuss some aspect of the thesis. However, it was very clear that they believed that it was their responsibility to make an independent judgement of the thesis: 'Examining a PhD is not about consensus' (Science/Male/2), and 'It's not what you are meant to do, to discuss amongst yourselves about whether to fail it or not' (Hum/Female/23).

Indeed, some experienced examiners regarded an approach from another examiner as an indication that the other examiner was unsure of their judgment. This attitude does raise problems for inexperienced examiners who might seek to consult their more experienced coexaminer: on the basis of this study, they would be better advised to consult an experienced colleague rather than the other examiner. However, a dissenting voice was raised on this point by one examiner who believed that consultation between the examiners was a better means of resolving disagreement than the common practice of sending the thesis to a third examiner: 'It would be better to encourage examiners to talk to each other, to exchange views on criteria and preliminary reports, and to give the student consistent advice' (Hum/Female/23).

Publications. The advice that is often given to postgraduate students is to publish as much as possible during their candidature. One reason given for this strategy is that when their thesis is presented for examination the examiner will be favourably influenced by the fact that the work has already been subject to peer review and found acceptable. Hence, we were interested to see whether experienced examiners were influenced by pre-publication of parts of the thesis. There were several points of view.

- Half our sample explicitly acknowledged that they were favourably influenced by the fact that a candidate's work had been accepted for publication in a reputable journal.
- For most of the remainder of the sample, in all cases examiners from the humanities and social sciences, the question was not particularly relevant to their experience, since pre-publication was not common practice in their disciplines.
- A small number of examiners expressed reservations about pre-publication, either because they were sceptical of the standards of many journals, or because they were concerned that the early publications might be the work of the supervisor or other members of the research team, whereas the thesis was a better reflection of the student's own work, or because the absence of publications might be due to a busy or negligent supervisor and not the student's fault, or because some acceptable thesis work, e.g. of an exploratory nature, might not lend itself to publication.
- A small number rejected outright the influence of pre-publications: 'The examiner is there to examine that piece of work, not anything else that they might have done' (SocSc/Female/3), and 'The examiner should be confident in his own opinion of the work and is not influenced by what peers might have said regarding published papers' (Sc/Male/22).

For those examiners acknowledging the influence of pre-publication, the extent of the influence varied. Some viewed publications as an 'insurance policy' (Sc/Male/9), or 'a check as well as a tick' (Hum/Male/17); that is, the examiner's judgement is being put beside other, very competent judgements, but this also 'lightens the burden for the examiner as other reviewers have said that it is OK' (Maths-Eng/Male/21). Others were much more forthright: 'It immediately suggests the student deserves the degree' (Sc/Male/26), and 'If there are two or three good publications you can put your feet up and go for an interesting drive. If there is nothing published you think "That's interesting"!' (Sc/Male/30).

Supervisors/departments. Of all the topics canvassed in these interviews, the one creating the most ambivalence amongst the experienced examiners was the influence of the student's supervisor and/or department on the examination process. As one said, 'The evaluation of a thesis is not a neutral event' (Sc/Male/10). This was also the only set of responses where we felt that some interviewees might not have been completely frank, in the sense that they felt that they were expected not to be influenced but they realised that they were influenced, at least to some extent. Tinkler & Jackson (2000 p. 172), in their discussion of the requirement that the examiner be 'independent', note how problematic such a characteristic is in an academic environment characterised by networks, membership of discipline-based academic bodies, previous supervisor/student relationships and collaborative research.

Some examiners were adamant that they were not influenced by the supervisor or department: 'You are marking the work not the supervisor' (SocSc/Male/15). Others conceded that knowing the supervisor influences the examiner's expectations about the approach that might be taken in the thesis rather than the quality: 'You might know that work is likely to go down a particular track because of the supervisor' (Hum/Male/8).

The influence of the candidate's supervisor or department on the examiners' expectations of quality operates in a twofold way: several examiners concede that they expect theses coming out of a highly regarded department to be of high quality: 'This one is going to be fine because I know the supervisor. I haven't even read the thesis but I know it will be OK—or I expect that it will' (Maths–Eng/Female/18). However, if the thesis is not as good as the examiner expects, the bias then runs against the student, both because there are few excuses for a poor thesis coming from a good department, and because examiners are disconcerted that a poor thesis has been 'allowed out' of such a department: 'Second-rate theses simply wouldn't come out of *that* lab' (Sc/Male/13). On the other hand, examiners are particularly pleased to receive a very good thesis from a little known department, or from a department or university that they know has fewer resources and facilities than the top class universities. In this latter case, examiners seem to be more confident that the good work reflects the student's talent rather than the supervisor's—when theses come from a top class department, there is always the concern that the student may have been 'carried' by the supervisor and his or her colleagues.

It should be pointed out that with a sample of experienced examiners, who were high-profile people in their disciplines, two factors come into play: they are unlikely to be sent a poor thesis that would reflect badly on the supervisor and/or department; and they usually know the supervisor who is sending them a thesis, because they know the other (high-profile) people in their discipline. As one examiner said, 'Everyone knows that in choosing examiners people don't choose examiners who will have all their buttons pushed' (Sc/Female/14).

Discussion

While it is not yet possible to fully develop a model of postgraduate assessment, the following components can be described. Experienced examiners make judgements about the quality and quantity of students' work based on their own extensive experience or mastery of the area. These examiners 'know' what constitutes a passable or outstanding thesis. They are also, on the whole, making judgements about the students' capacity, as demonstrated by the thesis, to undertake independent research. Indeed, experienced examiners make judgements about the quality of a thesis by the time they have read the first two chapters, often sooner. The positive indicators that lead to these initial judgements include:

- 'sparkle, élan and sense of confidence with the material' (Hum/Male/8);

- cohesiveness and clarity (Sc/Female/14);
- a student who makes the ideas his/her own, with some originality of presentation (Hum/Male/17);
- professionalism—as demonstrated by mature comments, and the accuracy of the logic (Maths–Eng/Female/20);
- style and sophistication (Hum/Male/17).

The negative indicators are:

- references that are poor: ‘This is usually a sign of a poor thesis—the two go hand in hand’ (Hum/Male/8);
- ‘Irritating things in the thesis such as typos and other careless textual mistakes that indicate a lack of attention to detail ... Sloppiness in the text indicates sloppy research’ (Sc/Female/14).

The final, substantive judgement is determined by:

- the student’s confidence and independence;
- a creative view of the topic;
- the structure of the argument;
- the coherence of theoretical and methodological perspectives; and
- evidence of critical self-assessment by the student.

In seeking to make sense of what experienced examiners report on their approach to the assessment of theses, we need to address several questions. Firstly, if examiners expect the thesis to pass, why have an assessment process? Experienced examiners take a holistic approach to the process of making judgements, in the way they consider the whole thesis document and judge the quality of its various aspects as they relate to one another, rather than as stand-alone qualities. Therefore, while it was possible to list a number of characteristics of a passable thesis earlier in this article, it seems that it is not possible to ‘mark’ each one out of 10, total the results, and declare a thesis passed or failed. The differentiating factor in terms of the purpose of the assessment is between what the student thinks the assessment is for and what the examiner believes it to be. The student sees the examination as summative assessment—passing, failing or grading. The examiner, on the other hand, sees the examination more as formative assessment, i.e. an exercise in giving feedback in an effort to assist the student in further developing and improving the work. It would be reasonable to suggest that many university staff and students consider that the thesis examination process is about deciding whether a student should pass or fail, i.e. be awarded or not awarded the degree. This approach to assessment, referred to as summative assessment, is similar to that of the undergraduate end-of-semester examination. While experienced examiners do invoke their ‘summative’ powers, it is clear that most experienced examiners approach the task from a formative assessment perspective. One examiner suggested that ‘The purpose of examining is to bring the work/the student up to speed’ (Sc/Female/12). Another examiner bemoaned the fact that most students were unable to see the process as a formative one; that is, ‘an opportunity for the students to be able to incorporate comments so that it sits on the library shelf and glows more brightly’ (Sc/Male/4). Arguing for the introduction of vivas into the examination process (currently only used in exceptional circumstances in most Australian universities), this same examiner commented, ‘I think that the oral as a means of giving the student feedback is an essential part of good teaching, but it is almost impossible to get students to see it as anything other than an examination’.

When examiners are in the process of examining, do they consider they are examining the thesis, as a piece of work, or are they examining the student? For example, do the

examiners base their judgements on the work as it stands in much the same way as they would an article for peer review, or are they using the thesis as a sample of the student's potential or capacity for independent research? This concept of 'capacity' differentiates the form of assessment undertaken at the postgraduate research level, compared with the undergraduate and/or coursework level. At the postgraduate level, the examiner seeks evidence that the student has demonstrated the potential to be an independent researcher, not that she/he is necessarily one already. Examiners seek evidence that the student has the repertoire of technical, intellectual and personal skills necessary to identify and tackle a 'do-able' problem (Fujimura, 1997).

For elaboration of this concept we asked examiners to talk about their views on the usefulness of oral examinations. We received two distinctly different answers, reflecting, we believe, two quite different views of what they were examining. One view was that it is the thesis, as a complete and comprehensive document that will remain on the library shelf, that is being examined. The other argument put forward was that it is the student as a potential researcher who is being examined and, therefore, ambiguities and perceived, although not necessarily demonstrated, potential should be teased out and followed up in a discussion with the student.

One discipline difference stood out in the responses, and that is that all of the interviewees designated as humanities believed that 'It is the text that is being examined, not so much what the students can argue and clarify in an oral' (Hum/Male/8). On the other hand, all of the mathematics/engineering interviewees and the bulk of the science and social science interviewees commented that, where there were ambiguities or uncertainties, it was useful to be able to clarify them with the student. This strong emphasis from the sample of experienced examiners indicates that, other than in the humanities, there is a belief that it is the student and her/his potential that is being examined, not the thesis document.

Based on the findings reported above, we are now in a position to be able to provide tentative advice to students and supervisors regarding the nature of the thesis examination process. We are also able to suggest strategies for inexperienced examiners and to outline policy implications for institutions.

For students, the most heartening information is that experienced examiners want them to be awarded the PhD and will go to extraordinary lengths to enable this to happen. The other information is that experienced examiners should be sought for the examination process, not avoided, because of their high degree of tolerance. These two factors arise from the examiners' experience as supervisors and their ability to judge the standard of a thesis based on a wide range of other examples. As one commented, 'As an examiner, you are not being a supervisor, but being aware of what students go through to get to that point makes one, hopefully, a wiser, less pedantic person and able to see what's being achieved. Also, you are able to see the vulnerability of the student' (SocSc/Female/7). As another said, 'I tend to be absolutely forensic when I mark a thesis and then I spend hours worrying about how harsh I've been ... have to argue myself into a sympathetic and tolerant framework. If you don't exercise tolerance it's very easy to mark a thesis' (SocSc/Male/15). Indeed, some interviewees expressed their reluctance to send their own students' theses to examiners outside the university system because people from industry or from research institutes might not understand the limitations of the postgraduate situation. It is feared that they might examine it as if the student has had several research assistants and a large grant to set up the research—in other words 'with little understanding of the student's situation' (Sc/Female/12).

Warnings to students are also clear from the research: careful attention to detail and the avoidance of sloppiness are essential. Sloppy presentation indicates to the examiner that the

research might well be sloppy. The other warning is the importance of being assiduous about actually doing what one says one is going to do, or explaining how and why changes have been made. The results of this research indicate that experienced examiners check carefully for the link between the introduction, where students outline their intentions, and the conclusions, where the intentions should have been realised.

Of course, supervisors might pass the above advice on to their students, and our experienced examiners had well-defined expectations about the responsibility of supervisors for the quality of the theses coming from their departments or laboratories. However, they did concede that students sometimes ignore the best advice of their supervisors and submit theses of inferior quality, yet their general view was that the quantity and quality of the work submitted reflected on the supervisor as well as on the student. They particularly held the supervisor responsible for helping students to deal with predictable difficulties, such as difficulties in writing in English. This research has demonstrated a clear link between experience as an examiner and experience as a supervisor. Examining a thesis helps to give supervisors a benchmark by which to evaluate their performance, and to more clearly appreciate their role as a supervisor. At the same time, experience as a supervisor gives one great insights into the examination process. Hence, for supervisors wishing to develop their skills and understandings, examining theses is seen as a very positive and helpful exercise.

For inexperienced examiners, the advice of their more experienced colleagues is unequivocal: 'Don't hesitate to seek advice'. In the older universities that advice is readily available and confidentiality rules should not prohibit examiners from seeking advice on specific matters. In universities with a smaller pool of experienced examiners to hand, more formal support programmes for novice examiners may be needed. However, there was little support among our interviewees for formal training courses for examiners. Their view seemed to be that the important discussion happens with respect to specific theses, so that generic professional development is not appropriate.

For heads of schools or departments or chairs of higher degree committees who choose examiners, the dilemma is how to appoint examiners expert in the thesis topic and avoid exposing their students to the shortcomings of inexperienced examiners. Anecdotal evidence indicates that some departments now make a point of avoiding inexperienced examiners. This practice raises the question of how novice examiners are to accumulate the necessary experience to be trusted as examiners. Based on suggestions from our interviewees and reflection on their views on other issues, we would recommend that:

- examiners' reports should be more widely available to potential examiners;
- confidentiality rules should not prohibit inexperienced examiners seeking advice from their more experienced colleagues;
- more formal mentoring programmes may be appropriate in some institutions.

Stage two of this research project, now being undertaken, may well help in understanding the specific issues related to inexperienced examiners. In this research we are interviewing a sample of inexperienced examiners who are also inexperienced supervisors, and a sample of inexperienced examiners who have had considerable experience in supervising.

However, on a much broader level, policy-makers and administrators have a challenge ahead if they hope to change the Australian PhD in any substantial way, as a result of changes to the way the Australian Government is funding research students' candidature and the time allocated to complete a PhD or Research Master's. Experienced examiners, as this research has demonstrated, 'know' what constitutes a passable or good thesis in terms of quality and amount. If theses are to be completed in shorter time periods, these examiners will need to carefully reconsider their views, and hence their judgements. This is particularly the case in

the humanities, where it is the 'product' which is examined rather than the demonstrated potential of the student to undertake independent research.

However, the final word of advice should go to students from one of our interviewees: 'A PhD is a stepping stone into a research career. All you need to do is to demonstrate your capacity for independent, critical thinking. That's all you need to do. A PhD is three years of solid work, not a Nobel Prize' (Maths-Eng/Female/18).

Acknowledgement

The authors gratefully acknowledge the generosity of the 30 examiners who contributed to this study through the interview process and by commenting on the draft paper arising from the interviews.

Correspondence: Dr Gerry Mullins, Adelaide Graduate Centre, University of Adelaide, ACT 5005, Australia; e-mail: gerry.mullins@adelaide.edu.au

REFERENCES

- BALLARD, B. (1996) *Contexts of judgment: an analysis of some assumptions identified in examiners' reports on 62 successful PhD theses*, paper to the Conference on Quality in Postgraduate Research, Adelaide.
- BOGDAN, R. & TAYLOR, S. (1975) *Introduction to Qualitative Research Methods: a phenomenological approach to the social sciences* (New York, Wiley & Interscience).
- BRENNER, M. (1985) Intensive interviewing, in: M. BRENNER, J. BROWN & D. CANTER (Eds) *Research Interview: uses and approaches* (London, Academic Press).
- BRUCE, C.S. (1994) Research students' early experiences of the dissertation literature review, *Studies in Higher Education*, 19, pp. 217-229.
- DELAMONT, S., ATKINSON, P. & PARRY, O. (2000) *The Doctoral Experience: success and failure in graduate school* (London, Falmer Press).
- DENZIN, N. (1994) The arts and politics of interpretation, in: N. DENZIN & Y. LINCOLN (Eds) *Handbook of Qualitative Research*, pp. 500-515 (Thousand Oaks, CA, Sage).
- FONTANA, A. & FREY, J. (1994) Interviewing: the art of science, in: N. DENZIN & Y. LINCOLN (Eds) *Handbook of Qualitative Research*, pp. 361-376 (Thousand Oaks, CA, Sage).
- FUJIMURA, J. (1997) *Crafting Science* (Cambridge, MA, Harvard University Press).
- HANSFORD, B.C. & MAXWELL, T.W. (1993) A masters degree programme: structural components and examiners' comments, *Higher Education Research and Development*, 12, pp. 171-187.
- JOHANSSON, B., MARTON, F. & SVENSSON, L. (1985) An approach to describing learning as change between qualitatively different conceptions, in: L. WEST & A. PINES (Eds) *Cognitive Structure and Conceptual Change*, pp. 233-257 (Orlando, FL, Academic Press).
- JOHNSTON, S. (1997) Examining the examiners: an analysis of examiners' reports on doctoral theses, *Studies in Higher Education*, 22, pp. 333-347.
- MILES, M. & HUBERMAN, A.M. (1984) *Qualitative Data Analysis: a sourcebook of new methods* (Newbury Park, CA, Sage).
- NIGHTINGALE, P. (1984) Examination of research theses, *Higher Education Research and Development*, 3, pp. 137-150.
- PHILLIPS, E.M. (1992) The concept of quality in the PhD, in: D.J. CULLEN (Ed.) *Quality in PhD Education*, pp. 11-21 (Canberra, Centre for Educational Development and Methods, Australian National University).
- PITKETHLY, A. & PROSSER, M. (1995) Examiners' comments on the international context of PhD theses, in: C. MCNAUGHT & K. BEATTIE (Eds) *Research into Higher Education: dilemmas, directions and diversions*, pp. 129-136 (Melbourne, Higher Education Research and Development Society of Australasia Victoria).
- TINKLER, P. & JACKSON, C. (2000) Examining the doctorate: institutional policy and the PhD examination process in Britain, *Studies in Higher Education*, 25, pp. 167-180.
- WINTER, R., GRIFFITHS, M. & GREEN, K. (2000) The 'academic' qualities of practice: what are the criteria for a practice-based PhD?, *Studies in Higher Education*, 25, pp. 25-37.

