

## PHD EXAMINATION AND EXAMINER CHARACTERISTICS

Sid Bourke, Jill Scevak & Robert Cantwell

SORTI, Faculty of Education and Arts

The University of Newcastle, Australia

### **Abstract:**

*The choice of examiners for PhD theses across disciplines tend to be based on a stable group of features, i.e. reputation in the field, publication record, subject and/or methodological expertise, and experience of research supervision and examination. In Australia each thesis is normally examined by three examiners, and there is the expectation that most or all examiners would be drawn from outside the examining institution, with possibly some from overseas. What are the results of this process? It is a question that is rarely posed, and yet may have a significant impact on the process of assessment. For any one thesis how 'balanced' are examiner characteristics and is there a difference by discipline? Do individual or collective examiner characteristics predict the ratings given to a thesis, or particular qualities of examiner reports? This paper draws on data from a study of the examination of 101 PhD theses from one institution (i.e. a total of 303 examiner reports). The findings have implications for the process of thesis examination, the interpretation of thesis reports, and advice given to supervisors and candidates.*

---

### RESEARCH TRAINING

The competitive pressures on university education and research worldwide are well documented. An increasing emphasis is being placed on maximising the skill and knowledge required for quality and impact of research in all areas of human activity, that is for all disciplines across the university. This emphasis extends to interest in all aspects of research training including student choice of research topic, supervision, progress, completion and satisfaction.

Completion statistics are one way to monitor research degree outcomes, student evaluations and examination success rates are others - and all three are controversial. Completion rate statistics underplay the complexities of candidature (Wright & Cochrane 2000). Student evaluation is built on a framework that is fundamentally atheoretical because sound evaluation procedures, although developing, do not exist for this specialised level of educational practice. Finally, thesis examination remains 'shrouded in mystery' (Johnston 1997, Tinkler & Jackson 2000). It is doubtful that work on the fundamentals of doctoral study has advanced sufficiently to alleviate concern expressed almost 15 years ago:

The absence of a research-based literature on doctoral study may have contributed to the apparent uncertainty about the nature, form and purpose of the degree. The purpose (or purposes) of the PhD have not been set down in such a way as would attract unequivocal and widespread agreement. (ESRC 1987)

Thesis examination is characterised by vagueness in what is looked for by examiners, partly caused by vagueness in specification of examination criteria. Tinkler and Jackson (2000)

found that the only specification that was common to all institutions was the requirement for the thesis to be an 'original contribution'. Although most institutions also specified that the thesis had to be located in the appropriate field and had to be the candidate's own work, this hardly constitutes clear guidelines for examiners. A more detailed and wide-ranging discussion of research in the general area of research training, and more specifically on thesis examination, has been provided in an earlier paper in this seminar (Holbrook 2001). The remainder of the brief review here concentrates on research thesis examiner selection.

## EXAMINER SELECTION

Perhaps not surprisingly given the above, questions have been raised about examination quality, the reasons given for examiner selection and their independence (Hansford & Maxwell 1993, Johnston 1997, Tinkler & Jackson 2000). Examiner selection assumes critical importance when, as Sloboda and Newstead (1997) indicate, that there are few, if any, checks and balances of examiners following proper process, and examiners may be relatively inexperienced. Sloboda and Newstead see these as two of several problems they see as associated with PhD examination.

From responses to a survey of 20 universities in the UK in 1999, Tinkler and Jackson (2000, pp.170-1) go so far as to suggest three criteria for examiner selection are the norm. The criteria are: (1) academic credentials - the examiner must be competent in the area and a specialist in the field of the candidate's research; (2) experience - at least one of the examiners should have examined at least three theses previously; and (3) independence - the external examiner should have had no formal attachment for a number of years, should not be used too often by the university, and under what conditions the candidate's supervisor can be an examiner. The findings are of considerable interest, although not entirely relevant to the Australian system where most, if not all, examiners are external to the university, and it is rare that a candidate's supervisor could also be an examiner. At the university involved in the present study (The University of Newcastle, Australia), all three examiners would normally be external to the university, with a special case having to be made for an internal examiner. In no circumstances could the candidate's supervisor be an examiner.

Criticism has been expressed of the emphasis on credentialing in examiner selection in tertiary education generally, and a consequent lack of public knowledge being involved in examination (Rawson 2000). The same writer would seem to favour an element of peer or self assessment, which he states are virtually non-existent. Even if this argument is accepted at other levels of university education to move assessment away from a too tightly-focused curriculum, given that the PhD degree is at the boundary of knowledge, it would seem necessary to select PhD examiners for their expertise. The PhD is different in kind as well as level from other tertiary studies - as an original contribution to knowledge a thesis does not have a re-producing element, and is not syllabus bound.

One final point about examiner selection was raised by Woods (1998) in relation to qualitative research methodology, but the point has a more general application. Woods was concerned that a narrow, perhaps traditional, methodological focus of an examiner could lead to discussion and dissent among examiners about the acceptability of the method to the ultimate detriment of the candidate. The general point is that an additional criterion for examiner selection is that they should be flexible in approach, and not only with respect to method. Examiners should be capable, professionally and personally, of recognising good work, even if the study differs from how they would have done it. This would suggest a strong familiarity with personal characteristics of examiners in addition to a familiarity with their work. The alternative is a sad one - the effective division of disciplines into schools of thought, with examiners selected by their membership of a school. This would become

restricting for examiners who would tend not to be asked to read 'different' theses, and less challenging for candidates who would no longer have a need to justify their methods.

Clearly there is evidence of concern over time for the process and practice of the selection of examiners of research theses. This study directly addresses some of these issues, specifically those associated with the reasons given for selection. The questions being asked are - what reasons or justifications are given for examiner selection, and are the stated reasons related to what the examiner does as assessed by the content of examination reports?

## UNIVERSITY REQUIREMENTS AND DATA COLLECTION

Thesis supervisors and heads of departments (or equivalent) in which the candidate has been enrolled are asked to provide the University Research Higher Degrees Committee (RHDC) with brief reasons for the selection of each of three examiners of each PhD thesis. The suggested reasons relate to examiner qualifications, expertise and experience. The examiner nomination form that departments are required to complete states:

... it would assist ... if the recommendation for the appointment of examiners was accompanied by an explanatory note concerning the appropriateness of the persons nominated, detailing qualifications, expertise in the field and previous supervisory and/or examining experience.

The Committee is also concerned to ensure the independence of examiners, expressed on the form through restrictions on the appointment of an internal examiner, and of examiners from the same department in any institution. To quote again from the nomination form:

The appointment of an examiner within the same department as the candidate is strongly discouraged and the appointment of more than one examiner from the same department at any institution is also strongly discouraged. It is acknowledged, however, that in exceptional circumstances, the expertise to examine may occur in such situations. In such cases a clear case must be made.

Going further than this with respect to examiners being independent, advice to each examiner of the identities of the other two examiners of a thesis is not provided by this university until the examination is complete and the candidate's result has been determined. This is not a uniform practice across Australian universities, but is by far the most common.

One other recommendation and warning is included on the nomination form to ensure some knowledge within the examination panel of Australian university research degree requirements.

It is important that, as a panel, the examiners have some familiarity with what is required by Australian universities for the award of Masters or PhD degrees. This might not be present if, for example, all of the examiners came from research institutes in non-English speaking countries.

It is rare for a nominated examiner not to be accepted by the Committee. However, there is sometimes discussion between the head of department and the chair of the Committee regarding some aspect of suitability, perhaps related to independence of a nominated examiner.

The most recent 100 thesis examinations (as at the end of 2000) across all disciplines at the university were the target sample for this study. Copies of the one-paragraph statements

recommending each examiner that heads of departments provide to the Research Higher Degrees Office for approval by the Chair of the RHDC, the Deputy Vice-Chancellor (Research), were collected by staff of the Research Branch, the legal custodians of the data. All identifying information was removed by the Branch before the statements were passed to the SORTI research team for coding and analysis. These statements were then coded for the 303 examiners (three nominations for each of 101 candidatures) involved in this data collection.

Two different types of approaches to examiner selection can be identified. The first was to nominate three examiners who will provide independent judgements of the thesis, but who would examine the thesis from the same, or at least similar, perspectives. The second was to nominate examiners with a view to them carrying out somewhat different functions. For example, one examiner might be a subject matter expert, another a methodologist, and the third might have recent experimental experience in a similar area. These two approaches are allowed for in the data collection, as analyses can be undertaken at the level of candidate (thesis) or examiner.

## REASONS GIVEN FOR EXAMINER SELECTION

Four major criteria - expertise, reputation, publications and experience - each having at least three specific components, were identified from the 303 statements provided by the heads of departments in nominating examiners. Recall that two of these criteria, namely expertise and experience, are specifically mentioned in the guidelines for examiner recommendation sent to departmental heads. The four criteria are briefly defined now in terms of their components.

### **Expertise**

Three main types of expertise could be identified. These were: (1) theoretical and/or empirical expertise, it being difficult to distinguish between these in the statements provided; (2) methodological and/or statistical expertise; and (3) general expertise, often expressed as having a strong interest in the area.

### **Reputation**

Six criteria of repute were distinguished: (1) international; (2) national; (3) general, often stated as 'highly regarded' in the field; (4) journal editor or president of a professional association; (5) member of editorial board or a leadership role in a professional association; and (6) honours or awards received.

### **Publications**

Three groups of statements about publication were developed: (1) international journals, exhibitions or patents; (2) national journals, exhibitions or patents; and (3) prominent publications generally.

### **Experience**

Three types of experience were given: (1) experience in PhD supervision; (2) experience in PhD examination; and (3) experience in the Australian university PhD system.

When one of these criteria was satisfied in a nominating statement, this was recorded. In some cases nominating statements included several statements that individually qualified as

fulfilling one of the criteria listed above. In these cases, the criterion was simply noted as having been satisfied. Multiple instances of a particular criterion were not recorded. Consequently the maximum number of criteria that could be scored for any one examiner was 15 codes.

There are at least two ways of considering the breadth and focus of the reasons given for examiner selection. First there is the set of reasons stated for the selection of an individual examiner. The reasons an individual examiner was selected are likely to be important for the focus of their examination and hence the reports they write, and the nature and content of examiner reports are a major focus of this study. Consequently the criteria stated for individual examiners will be described. Secondly the reasons given for the panel of three examiners selected for each candidate's thesis can be considered collectively. This approach is more likely to be important when considering the consistencies and differences in the nature of the reports written and the comments made by individual examiners about the same thesis. The results of both of these approaches are now described, commencing with the reasons given for individual examiner selection.

#### INDIVIDUAL EXAMINER SELECTION

Reasons given for individual examiner selection are described in two ways. First the frequency of inclusion of the specific criteria for the whole sample of 303 examiners are shown as percentages, grouped according to the general criteria in descending order of occurrence (see Table 1). Secondly, the general criteria are described by the five major Broad Fields of Study (BFsOS) which had substantial numbers of these included in this study, with differences in relative frequency by BFOS being noted (Table 2).

As it was included as a reason for selection of more than two-thirds of examiners, theoretical and/or empirical expertise was by far the most common specific criterion given. It was followed by publications generally included for more than one-third. International and general reputation and professional leadership, and experience in supervision and examination were also common criteria given. When the specific criteria were grouped, expertise was by far the most frequently-given general criterion, being included for more than three-quarters of the examiners. Reputation was the next most frequent general criterion, followed by publications and, finally, experience, cited as a reason for more than one-third of the examiners.

There are some clear differences between fields of study on all three of the four general criteria, the exception being expertise. Differences between BFOS are statistically significant for the other three criteria. In Table 2, the criteria are again taken in descending order of frequency overall. To provide continuity, the last line of Table 2 is the same as the first column of Table 1.

TABLE 1. EXAMINER SELECTION: REASONS GIVEN

<b>GENERAL CRITERIA</b>	<b>% of Total</b>	<b>SPECIFIC CRITERIA</b>	<b>% of Total</b>
EXPERTISE	77	Theoretical/Empirical	68
		General (eg, strong interest in area)	10
		Methodological/Statistical	8

REPUTATION	59	International	28
		General (eg, highly regarded)	17
		National	13
		Editor of Journal/President of Assoc.	12
		Member of Edit. Board/Leader. Assoc.	7
		Honours or Awards	3
PUBLICATIONS	40	Publication generally	35
		International journal/exhibition/patent	6
		National journal/exhibition/patent	2
EXPERIENCE	36	In PhD supervision	23
		In PhD examination	20
		In Australian university PhD system	8

Although expertise was the most frequently given reason for selection of all examiners, it was clearly higher for Science (88%) than the other BFOS. Health was somewhat lower than the other BFOS, although expertise was still a criterion for selection of more than two-thirds even there. Reputation was far more frequently stated as a selection criterion for Engineering more than for the other BFOS, and reputation was the least mentioned for Health and for Science, although still for almost half the examiners in both fields. Arts and Education more often gave publications as a reason and Science gave it less often than the other BFOS. Science was particularly high on experience as a reason given for selection of more than half the examiners, compared with the other BFOS, particularly Arts and Health with one-quarter or less.

TABLE 2. REASONS GIVEN FOR INDIVIDUAL EXAMINER  
SELECTION BY BFOS

REASONS GIVEN x BFOS	EXPERTISE %	REPUTATION %	PUBLICATION %	EXPERIENCE %
ARTS	75	60	55	25
EDUCATION	76	57	54	31
ENGINEERING	78	82	35	37

HEALTH	68	47	37	23
SCIENCE	88	48	25	55
ALL BFOS	77	59	40	36

### SELECTION OF THE EXAMINATION PANEL

In conjunction with individual examiner selection, the composition of the panel of three examiners is important in ensuring all major aspects of the thesis being examined are covered. The approach taken here to describe the collective reasons given for the panel selected to examine a thesis was to indicate the percentage of theses where each of the four general criteria was included as a reason for selecting at least one of the examining panel. The first thing to be noted was that the order of frequency of the general criteria for the panel was the same as for the examiners taken individually (see Table 3).

The most common general criterion given for the selection of at least one of the examining panel for each thesis was expertise. Conversely, however, expertise was not given as a reason for selecting any of the panel for 9% of the theses. Across the major BFOS in this study, almost all of the Arts and Science panels included at least one examiner selected for expertise. At the other extreme, as 'few' as 85% of Engineering theses had one or more of the examination panel selected for expertise.

Reputation was the second most frequent general criterion given for examiner panel selection across all BFOS, with 83% of panels including this reason. There was no Engineering candidate who did not have at least one examiner on the panel selected for this reason. Even the BFOS which gave this reason least, that is Health, still had 70% of its panels including reputation.

Publication was next in frequency of examiner panel selection, and 61% of panels across all BFOS included this criterion. Consequently publication was well below the frequencies of expertise and reputation as a criterion. However, there were still about three-quarters of the Arts and Education theses that included publication in the examiner panel selection, while this reason was included in about half of the Engineering and Science panels.

Experience was given as an examination panel criterion for 57% of the theses examined across BFOS. Experience was included in less than half of the Arts and Health panels, but 70% of the Science candidates had at least one examiner selected for experience on their panel.

TABLE 3. PROPORTION OF THESES BY BFOS WITH AT LEAST ONE OF THE PANEL OF EXAMINERS SELECTED ON EACH OF FOUR CRITERIA

REASON GIVEN FOR CHOICE / AVE. OF BFOS (%)	REASON LESS OFTEN GIVEN / BFOS (%)	REASON GIVEN ABOUT AVERAGE / BFOS (%)	REASON MORE OFTEN GIVEN / BFOS (%)
--	------------------------------------	---------------------------------------	------------------------------------

EXPERTISE (91)	Engineering (85)	Education (89) Health (90)	Arts (95) Science (95)
REPUTATION (83)	Health (70)	Arts (80) Education (84) Science (85)	Engineering (100)
PUBLICATION (61)	Engineering (50) Science (45)	Health (65)	Arts (80) Education (74)
EXPERIENCE (57)	Arts (40) Health (40)	Education (58) Engineering (50)	Science (70)

Two points of interest arise from these results. First, although both expertise and reputation were included in the examination panels for almost all theses, there were differences in the relativities between BFOS. Thus Engineering was in the highest category for reputation and the lowest category for expertise, although in both cases the proportion is quite high. For the less frequent reasons, publication was given for most Arts panels while experience was given for less than half the panels in the same field. Secondly, different examination cultures would seem to operate across the major BFOS at this university. It could be claimed that Arts panels were selected largely on the basis of expertise and publication, Engineering mainly on reputation, Science on expertise and experience, and Education on publication. Health panels had no dominant reason although these shared in the consistently strong emphasis on expertise across all areas.

#### SELECTION AND EXAMINER COMMENTS

There are some clear differences between the frequency of reasons given for examiner selection across four main criteria identified, that is, expertise, reputation, publication and experience. Although many examiners were selected on more than one of these criteria (and two examiners were identified as having all four criteria), some were selected on only one of the identified criteria. There was also a small group of five examiners where the reasons for their selection could not be classified as any one of the four criteria. In these cases some reference was generally made to personal contact with the examiner being nominated.

The two most prevalent reasons given for selection are used to illustrate the pattern of some of these selections. A total of 32 examiners were selected for expertise alone, and 69 were selected on criteria not including expertise. And nine examiners were selected only on reputation with 123 selected on criteria not including reputation. Because they might provide the greatest contrast in the text content of their reports, two interesting groups to compare would be those selected only for expertise (n=32) and those selected only by reputation (n=9).

The content of the examiner reports was coded according to five parent nodes, each with several subsidiary nodes. The parent nodes are sectioning, examiner and process, assessable areas, dialogic elements and evaluative comments. These and the more detailed coding categories have been described in some detail elsewhere (see Holbrook, Bourke,

Farley & Carmichael 2001) and are appended to another paper being presented in this seminar (Holbrook, Lovat & Hazel 2001).

Several relationships between reasons for selection and the content of examiner reports could be expected. Examples from each of the four selection criteria will be used to illustrate expectations. Examiners selected for expertise would be expected to write more on assessment of the research approach taken in the thesis. Those selected on reputation might place more emphasis on formative and less on prescriptive instructional comments. Selection on publication record might suggest a stronger contribution relating to publications that might arise from the thesis. Finally, selection on experience as a supervisor might lead to more editorial comments, while experience as an examiner might lead to more commentary on the process and criteria for PhD theses. These and other relationships were tested.

Generally speaking, there were no significant patterns of correlation between reasons for examiner selection and specific content of examination reports. Most of the correlation coefficients anticipated as being significant were close to zero. Table 4, however, shows the mean text units (lines) for four content codes as a proportion of the complete reports written by the two groups of examiners referred to above as most likely to be different, i.e. those examiners selected solely for expertise and for reputation. Only the means for those codes where there was a statistically significant difference (at the  $p < .05$  level) between expertise and reputation are shown, together with the mean numbers of text units for the two groups to provide a scale against which to judge the percentages shown for each code. A t-test for independent samples was used to test the mean differences, making an appropriate adjustment in those comparisons where the group variances exhibited significant heteroscedasticity.

TABLE 4. CONTENT OF EXAMINER REPORTS BY SELECTION FOR EXPERTISE AND REPUTATION

CONTENT CODE	EXPERTISE MEAN % (SD)	REPUTATION MEAN % (SD)	PROB. LEVEL
Assessment: inaccuracy in review of literature	2.6 (6.7)	0 (0)	.036
Dialogic element: conversational	8.5 (14.1)	2.6 (3.4)	.036
Evaluative element: summative neutral comment	2.7 (3.6)	0 (0)	.000
Evaluative element: other instructional prescription	10.3 (12.8)	0.4 (0.8)	.000
Number of text units in report	N = 92 (70)	N = 109 (156)	NS

Examiners selected solely for reputation did not comment on any inaccuracies in the literature review and did not make any neutral summative comments on the thesis. In general neutral comments have been interpreted fence-sitting at best or as mild criticism at worst. In contrast, those selected solely for expertise gave about 2.5% of their reports to each of these types of comment. As reputation would seem to be a more general criterion than expertise, perhaps those selected on reputation did not concern themselves with what could be

seen as detail and unnecessary comment. In a similar way, examiners selected for expertise gave more than three times the emphasis to conversational elements in their reports compared with examiners selected for reputation. Finally, and most noticeably, examiners selected for expertise gave many times more emphasis to prescription (more than 10% of their reports) indicating in no uncertain terms that elements of the thesis required 'fixing', than examiners selected for reputation who did this to a minimal extent (less than 1%).

## CONCLUSIONS

It is notable that the guidelines provided by the university highlight expertise and experience as requirements for examiner selection, but only about three-quarters and one-third (respectively) of the recommendations made included a statement that can be interpreted as meeting these aims. And these recommendations were accepted by the Chair of the committee setting the guidelines. This is not to say that the head of department making the recommendation did not believe these requirements were not met by their nominee, simply that he or she did not mention the reason in the statement provided. Although reputation was not explicitly mentioned in the same guideline, almost three-fifths of the examiner recommendations made did include this as a reason for selection.

There were also some indications that reasons for examiner selection may be, to some extent, discipline based. Although all the major BFOS included expertise in almost all their examination panels, there was much more variability between fields for the remaining three selection criteria identified, namely reputation, publication and experience. There may be interest in pursuing these differences in any future studies that compare examination practices across discipline areas.

When the reasons given for the individual selections of 303 examiners were considered in conjunction with the content of examiner reports, it was clear that there were no strong relationships between the reasons and content, and only a few statistically significant ones. This leads us to the conclusion that, even if it is the intent of the head of department that an examiner play a specific role in the examination process by focussing on particular aspects, the role actually played by the examiner generally bears no close relation to the role intended. The content of the examiner's report is generally not distinguishable from the content of other reports written by examiners selected for quite different reasons. Of course, it is also possible that the head of department did not really select the examiner for the specific reasons stated, but his/her intent was simply to meet a university requirement to provide such a statement.

We were led to undertake the work described in this paper by interest in PhD examination and, in particular, by concern expressed in the literature about one important aspect - criteria used in examiner selection. However, the analyses of reasons given and their relationships with examiner reports described in this paper would lead us to suggest that there is little to gain by collecting statements of the 'official' reasons put forward for examiner selection in an attempt to throw further light on other information gathered as part of a study of PhD candidature and examination. As we intend this to be an on-going study, the time and effort saved in subsequent stages of our investigation will be considerable. We would further suggest that any new or continuing work on examiner selection should attempt to find other methods of investigation of selection criteria for this significant aspect of PhD candidature and examination.

## REFERENCES

ESRC (1987). The Social Science PhD: the ESRC inquiry on submission rates (Winfield Report).

Hansford, B.C & Maxwell, T.W. (1993). A masters degree program: structural components and examiners' comments. *Higher Education Research and Development*, 12(2), 171-187.

Holbrook, A. (2001). PhD examination: The barely mapped frontier in assessment. Seminar paper presented at the Annual Conference of the AARE, Fremantle.

Holbrook, A., Bourke, S., Farley, P. & Carmichael, K. (2001). Analysing PhD examination reports and the links between PhD candidate history and examination outcomes: A methodology. *Research and Development in Higher Education*, 24, 51-61.

Holbrook, A., Lovat, T. & Hazel, G. (2001). What qualities are rare in examiner reports? Seminar paper presented at the Annual Conference of the AARE, Fremantle.

Johnston, S. (1997). Examining the examiners: an analysis of examiners' report on doctoral thesis. *Studies in Higher Education*, 22(3), 333-347.

Sloboda, J.A. & Newstead, S.E. (1997). Guidelines for PhD examiners: an evaluation of impact. *The Psychologist*, September, 407-410.

Rawson, M. (2000). Learning to learn: more than a skill set. *Studies in Higher Education*, 25(2), 225-238.

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.

Woods, B. (1998). Issues in the supervision and assessment of doctorate theses using qualitative approaches. *Clinical Psychology Forum*, 114, 28-29.

Wright, T. & Cochrane, R. (2000). Factors influencing successful submission of PhD theses. *Studies in Higher Education*, 25(2), 181-196.