Australian Journal of Educational & Developmental Psychology. Vol 4, 2004, pp 126-145

Qualities and Characteristics in the Written Reports of Doctoral Thesis Examiners

Allyson Holbrook[1], Sid Bourke, Terence Lovat, Kerry Dally
University of Newcastle

ABSTRACT

This paper outlines the procedures used in the textual analysis of examiner reports for 101 PhD candidates across disciplines in one Australian University. The method involves the use of QSR software[2]. Three levels of findings are outlined. The first level is the coding categories that emerged out of reading the report text. There are five broad categories of codes that capture: the structure of the reports, the ways in which examiners communicate, the subject matter of the thesis, the characteristics of examiners' evaluative comment and their comments on their role and the examination process. The second level of findings concerns the frequency of different categories of comment and the prevalence of comment on the analysis and interpretation of the candidate's results. The third extends beyond the individual categories to what we can learn about the utilization of the report. One key finding is that the examiners took on specific roles: mentor-colleague, supervisor-instructor and assessor-arbiter. It was also found that the examiner tends to "enter the examination" at a point beyond the proposal, that is they comment primarily on elements that they feel they can influence, and this influence was evident in the preponderance of the formative instruction provided.

Keywords: doctoral assessment, examiner reports, text analysis, thesis quality

The objectives for doctoral study are not articulated in conventional ways, which means that the learning that takes place at doctoral level is something of a mystery. Confusion about what constitutes appropriate doctoral research and the lack of clear standards for the thesis have already been identified as problems within the academic community (Noble 1994, 32-33, 39), but where to start? In Australia the research focus has tended toward supervision, but PhD examination also presents an avenue of investigation with considerable potential. Given that thousands of theses are examined annually, it is clear that examiners are applying criteria and standards of some kind. How these criteria are defined and how consistently they are applied across disciplines, institutions and nations remains to be explored.

In Australia, as in many countries, the completed research doctorate is presented in written form as a thesis (or dissertation) and examined externally but, unlike some countries, the thesis is the primary and, in many cases, the only evidence of the candidate's learning and skills development. There is usually no oral examination. Given that the thesis is the end product of a long process, what interpretative frameworks do examiners apply and are these evident in the reports? This particular paper draws on the core text analysis of examiner reports on 101 theses (see Holbrook & Bourke, this issue) from one university. The study is

[1] Contact
A/Prof Allyson Holbrook
School of Education
University of Newcastle
NSW, Australia
Email allyson.holbrook@newcastle.edu.au

[2] The reports were originally analysed using QSR software Version N5 but have subsequently been analysed using QSR N6.

also the first in a series of replication studies involving a total of eight Australian universities with various levels and areas of research strength.

The main research questions underpinning the analysis in this paper are:

- What types of evaluative and other comment are contained in the examination report?
- What qualities, attributes and characteristics of the thesis are examiners emphasizing in their reports, and in what ways?
- How do examiners make use of the report in the assessment process?

Tinkler and Jackson (2000, 2001) and Jackson and Tinkler (2001) investigated the examination process in England, obtaining documentation from 20 universities (based on a stratified sample of old and new universities). In addition, they drew on questionnaire responses from some 100 examiners and candidates from two of the 'old' institutions. In Australia, Mullins and Kiley (2002) collected data from a small number of institutions and 30 examiners. Johnston (1997) collected 51 examiner reports from one institution across five faculties over several years. Pitkethly and Prosser (1995) utilised the reports of 74 thesis candidates at one institution. There were some common threads in the findings including general agreement among examiners and institutions about the core principles involved, namely that the thesis demonstrate originality and make a contribution. Evidence from a comparative cross-national survey by Kouptsov (1994) further bears this out. However, the concepts are slippery and contextualised by processes and procedures. What does it mean to apply them? How are they judged? Such questions have never been adequately addressed. The complexities of process are highlighted by findings that examiners treat the thesis as an end in itself (Johnston, 1997; Pitkethley & Prosser, 1995), yet rarely recommend an outright failure. In their study involving interviews with experienced Australian examiners, Mullins and Kiley (2002) noted that examiners appeared very clear in the distinctions they made between poor, acceptable and outstanding theses, but they also detected that examiners went into the process anticipating that students would pass. In the 303 examiner reports that constitute the data reported in this issue, only two examiners recommended a grade of fail. It would generally appear that examiners are inherently interested in undertaking an examination and approach the task in a positive light (Johnston, 1997; Jackson & Tinkler, 2001; Mullins & Kiley, 2002). However, a poorly written thesis generally had a negative effect on the examiner, suggesting disengagement (Johnston, 1997; Mullins & Kiley, 2002). A panel of 67 scholars from the USA, UK, Australia and Canada identified writing quality as one of the most problematic issues about PhD study (Noble, 1994). Most researchers in the field have discovered that editorial errors and presentation issues attract a substantial proportion of examiner comment.

Institutions tend to differ in the guidelines they provide for examiners. Tinkler and Jackson (2000) identified considerable procedural variation in examination across English universities, as well as vagueness in terms used. On the basis of content analysis of reports, Johnston (1997) found examiners tended to follow university guidelines or recommendations about how to report on a thesis, whereas Mullins and Kiley (2002) reported the opposite on the basis of interview data. They found examiners had established their own criteria, and noted but did not use guidelines provided.

This article offers a more fully explicated textual analysis of examiner comment than has been reported in the literature to date. It defines the core coding categories that are the basis of the statistical analyses to be reported in subsequent publications, while advancing the findings to the questions stated above. The latter emerged through the process of discovering, naming, unpacking, comparing, verifying, refining and finally "counting" the text units in, those categories.

1. Coding procedures and category verification

The text analysis that is the subject of this paper represents the first phase of coding and analysis for the 'PhD examination project' (referred to hereafter as the 'core'). It is the core analysis that is translated into text unit counts and feeds into the statistical analysis that enables links to be made between the types of examiner comment and their overall

recommendation or rating for the thesis, and identifies whether examiners cover the same ground within and across disciplines. The coding journey has been a long one and is by no means over. Future cross-institutional phases will involve extended analysis of the text that will require fresh lines of inquiry, input by discipline specialists and the application of discourse analysis.

The reports (which are on average 3.5 pages long after standardization into a line format that allows accurate text count comparison) are idiosyncratic and richly textured. On the one hand the coding process had to capture this phenomenon including the full scope of the comment. On the other, the aim was comparison, which required a coding framework that could be duplicated by multiple coders across discipline areas and institutions. Although we had theorized this approach was possible, there was no guarantee the aims would be compatible, or that such a coding scheme could be developed from examiner reports. Previous published studies of examiner comment had been sparse on detail with respect to how coding consistency was achieved. Hence for this project there was an extended process of trial and error undertaken primarily by the first author but which drew on the project team at critical points.

The steps that have led to the finalisation of the core stage were as follows:

1. Development of temporary categories (based on examination experience, institutional guidelines and the literature) to provide some scaffolding for engagement with the text. The first level of intensive reading of the text occurred after scanning, during preparation for entry into QSR N5. N5 is text analysis software and provides for overlapping and hierarchical coding consistent with conceptual richness. A new set of coding categories (referred to as nodes when utilizing N5 software) emerged from this stage of reading that reflected the overt characteristics of the reports. The characteristics encompassed the content of the thesis and its qualities; features of the text that told us something about the examiner and the process they were engaged in; and organizational features of the report text. The top level or parent coding categories are referred to hereafter as the primary coding categories.
2. Twenty candidate cases (generating 60 examiner reports) were used to determine if a reasonable rate of agreement was emerging between coders on the primary categories. This led to the process of writing detailed coding notes, and identifying text examples from the reports as illustrations (a feature that is utilized subsequently in this article). As a result of this process sub-categories were also identified, but tentatively and in small numbers because the methodological literature on text analysis, and the experience of the authors warned against the unmanageable proliferation of categories.
3. Approximately one half of the reports from the first institutional case were then coded and, during this phase, hierarchical coding categories evolved and inter-rater agreement was tested. Ninety per cent agreement was taken to indicate stability of the category (see Holbrook & Bourke, this volume). Initial text unit counts and correlational analyses of the categories with the examiners level of recommendation provided early indications of examiner emphasis. The findings were used to interrogate the existing themes and advance sub-categorisation, for example, this early statistical analysis was used to indicate whether categories were predicting the rating given to theses in a commonsense way and alerted us to the possibility of combining some sub-categories, moving others to another parent altogether, and unpacking others into new sub-categories.
4. Tentative forays into the symbolic structures and discursive features of the text in conjunction with exploratory quantitative analyses led to the first major finding, namely the identification of features of report utilization. When we started to explore that avenue, and juxtapose our tentative hypotheses with the literature, the evidence mounted that examiners positioned themselves in various ways, taking on specific roles. Moreover, there were two areas of text that were relatively under-represented – comments on methodological approach and comments on the significance of the thesis. It was a productive moment in the analysis, because we were experiencing

difficulty teasing out the type of evaluative comment we had captured to that point. A new focus on examiner role took us back to the fundamentals of assessment and its purpose. We began to code examiner comment by its implication for action, identifying whether it was instructive or not and, if instructive, whether it was formative or not. This process resulted in three entirely new sub-categories of evaluative comment. We had discovered evaluative comment that was clearly formative, other comment that was entirely prescriptive, and a form of instructive commentary that was neither. We had also identified comments that were imbued with judgment but not intended to bring about any direct action. This activity produced five primary coding categories including: Report Organization; Examiner and Process; Assessable Areas Covered; Communicative Competence; Dialogic Elements; and Evaluative Elements.

5. Sections of the first half of the reports were then re-coded to include the new categories, and simultaneously the second half of the reports were coded from the beginning with particular attention to the continued refinement of the coding of the text specifically concerned with the thesis and its qualities. The authors remained alert to the possibility of entirely new information that was not addressed in the existing groupings. No such information was discovered, however, the range of examples for each category and sub-category increased providing more fine-tuned coding notes. Exploratory factor analysis was performed on the data from the first group, and confirmatory factor analysis was subsequently performed on the data from the second group. Consistencies between the two groups of data provided further evidence that we could successfully duplicate the coding process.
6. In QSR N5 text can be captured by category (chunking) for further analysis. Chunking makes it easier to detect coding inaccuracy. Moreover the impressions generated by this 'focus' can challenge the robustness of the coding framework in relation to the research questions, and the strength of the emerging findings. For the core analysis the final major check of the coding occurred when we returned to coding the positive and negative orientation of the evaluative comment (referred to as "tone" – see Holbrook & Bourke, this issue). This process in turn necessitated a major re-reading of a large proportion of the text. What this process mostly alerted us to, however, was the degree of overlap between the type of content covered and the type of judgment applied, also to the relatively small amount of overlap between the evaluative sub-categories which will in the future allow us to get a strong sense of the mix of positive and negative comment – an aspect we have already started to explore in relation to examiner comment on resubmitted theses (Lovat, Holbrook, Bourke, Dally & Hazel, 2002).

In summary, before we were ready to move to the next set of examiner reports from another university, we ensured that the categories we had developed were consistently being used, that our coded data base was free from avoidable errors (such as the occasional miscoded line), and that the categories themselves were inherently meaningful (albeit at a coarse-grained level). Furthermore, we ensured that the categories were inclusive of all the data in the reports, were immediately useful in guiding productive lines of questioning, and were theoretically informed.

The listing of categories and sub-categories that remained to be included in the final analysis is appended to this article. In QSR N5 coding categories are labeled and numbered, and the numbering indicates a hierarchy (see Holbrook & Bourke, 2004). In the appendix to this article some of the numeric indicators are not consecutive, but this is typical of a developing coding framework using N5. Another indication that this is a developing framework is the prevalence of the category numbered 10. In this project, category '10' signifies tentative or transitional categories or occasionally ones that serve a signposting purpose for later reference. It has proved useful to have one such category under each primary category to give a profile to 'other' text (for example, possibly unusual or anomalous text). In light of the discussion above, an example of the latter is the sub-category (5 10 2) that tags the rare occasions where thesis resubmission is called for. This example highlights that some

coding is primarily for the purpose of 'navigation'. Mostly, by the conclusion of the core-coding phase, the categories numbered "10" included no text because the material was already re-located to another sub-category. The attention to such detail is neither typical nor necessary in many text analysis projects, unless, as in this case, text counts need to be as accurate as possible for future reference, and all text needs to be accounted for.

2. Common elements in the reports

The five core primary coding categories are defined in Holbrook and Bourke (this issue). They are Report Organization (1), Examiner and Process (2), Assessable Areas Covered (3), Dialogic Elements (4), and Evaluative Elements (5). Each of the five categories captures distinctive features or characteristics of the report. The categories Report Organization and Dialogic Elements, however, can best be described as capturing features about report presentation and communication, including the manner of examiner engagement with the task of report writing. The other three categories capture features more specifically related to the knowledge base and expectations underpinning examination. Most of the text in examiner reports is about assessment and evaluation, so that is where the main discussion about the qualities and characteristics begins in this article. Following the section on assessment and evaluation is a brief discussion of the structural and personal features of the reports. In the final section, the questions raised earlier in the article concerning examiner role will be addressed.

2.1 Assessment: themes and emphases

Examiner attention ranges across a number of assessment areas which in this study are captured under core category (3) Assessable Areas Covered. There are no surprises in the areas identified, as they might be found in any guide to writing a thesis, and have been identified under similar headings in other research (Hansford & Maxwell, 1993; Johnston, 1997). The specific areas are "scope, significance and contribution," "review of the literature," "approach," "subject matter/findings," and "communicative competence." Some of these in turn have a further level of sub-categorisation (see appendix) and the latter occur in the table below slightly reworded from the coding labels to assist clarity of communication. Before turning to the results it needs to be recalled that the coding category definitions emerged out of the text and the meanings attached to the labels were similarly shaped. They do not necessarily reflect textbook definitions, which is why an elaboration on each of them can be found below. Table 1 provides the percentages of theses that included at least one instance of each sub-category in Assessable Areas Covered and the proportion of comment devoted to each sub-category as a percentage of the total report.

2.1.1 Scope, significance and contribution (3 1)

Scope (3 1 1) captures examiner reference to the candidate's research questions and aims. It is not unusual for an examiner to paraphrase or quote what the candidate claims the scope of the study to be. Less frequently examiners may elaborate on the questions and summarise what the study is about. Comments about the scope of the study occurred in 45 per cent of reports but, as indicated in the following example, were generally brief (constituting only 3 per cent of the total report).

> *This thesis describes the isolation and characterisation of secondary metabolites from four Australian brown seaweeds and a sponge.*

Comments about Contribution and Significance (3 1 2) were found in 70 per cent of reports. As with scope, such comments were typically brief (constituting 9 per cent of the total report) and frequently restricted to simple statements using common words such as 'useful', 'interesting' 'important', 'timely' and 'worthwhile.'

> *This is a very worthwhile study of an interesting and significant topic…It is notable for its substantive contribution to our knowledge and understanding of pupil's experience of school…*

Under such circumstances it is what examiners go on to say about the components of the thesis and their relative strengths that will give a sense of the contribution they identify. The relative paucity of elaborated comment on significance and contribution did prove a surprise

given the emphasis on these characteristics in doctoral study. The finding flags at least two questions. The first is: do examiners expect to comment at length on this feature, and the

Table 1: Assessable areas covered – occurrence of comment and proportion of total report (N=303)

Assessable areas sub-categories	Occurrence In reports [b] %	Proportion of comment [c] %
Scope (3 1 1)	45	3.4
Significance and contribution (3 1 2)	69	8.8
Publications arising (3 1 3)	27	1.3
Publications existing (3 1 4)	11	0.5
Literature – coverage (3 2 1)	53	4.0
Literature –error &/or inaccuracy (3 2 2)	23	1.5
Literature – utilization &/or theoretical application (3 2 3)	24	1.7
Approach (method & methodology) (3 3)	60	6.6
Analysis & reporting (3 4 1)	89	43.1
Topic related issues raised by the examiner (3 4 2)	44	10.4
Substantial issues of communicative competence (3 5 1)	73	5.4
Editorial comments (3 5 2)	50	8.3
Ethics (3 10 1)	2	0
Listing of what is in the thesis ('ingredients') (3 10 3)	5	0.3

[a] A full listing of every sub-category can be located in the appendix.
[b] The sub-category occurs at least once in a report.
[c] The proportion of comment from each sub-category as a percentage of the total report

second is: are they used to doing so? It may be that elaborating on the significance of the work is not a common feature of research evaluation.

Publications are a specific type of contribution and also flag significance if peer reviewed. Where publications are suggested or anticipated by the examiner, this is coded at (3 1 3). Where the examiner notes, reports or comments on publications that already exist, this is coded at (3 1 4). These two aspects of publication occurred in 27 and 11 per cent of reports respectively. Pitkethley and Prosser (1995) explored the extent to which examiner comments placed theses in an international context and one indicator that they identified was reference to publication:

> *Recommending the publication of aspects of the thesis was the most common means by which examiners linked the research to international, academic debate within the field. While recommending for publication is indeed the highest compliment, the fact that so few examiners gave suggestions where the candidate's research could move in the future, in a personal context, was surprising. (p 135)*

Pitkethly and Prosser (1995) reported that eight per cent of Australian examiners (N = 107) and ten per cent of international examiners (N = 120) discussed work already published. Our findings for 303 examiners (47 per cent of whom were international) were very similar, with between 10 and 11 per cent of both Australian and international examiners commenting on existing publications. Pitkethly and Prosser also reported on comments related to the potential for the candidate to publish. In this case they found a larger difference, with 27 per cent of Australian and 35 per cent of international examiners making such a reference. In the present study the proportions were reversed, with 31 per cent of Australian and 23 per cent of

international examiners commenting on possible publications arising from the candidate's project.

2.1.2 Review of literature (3 2)

Even though some form of a literature review is included in every PhD thesis, only about one half (53 per cent) of the examiners refer to it in their comments. Wherever the examiner draws attention to the literature, the comment is coded as sub-categories of (3 2). However, when a secondary analysis of the literature is the focus of the study or an integral part of method, then the pertinent text would be coded as methods and subject matter, that is, at categories (3 3) and (3 4).

In most documents the discussion of the literature by the examiner will be restricted to Coverage (3 2 1), that is the depth, breadth, recency, and adequacy of the literature used in the thesis.

> *The strength of this thesis is in the thoroughness of the literature review. Over 400 references are included in the study, with over 80% of the cited references having been published in the last 10 years. The review should be a very useful resource for other researchers in this field.*

If the examiner detects inaccuracies, the comment relating to that is coded at (3 2 2). Inaccuracy encompasses comments that refer to misreporting or inconsistent reporting of the literature, for example, in referencing, citation or quotation.

> *Far too many errors occurred in the References. At least 15 references were made in the text, but were not listed in the Reference section. At least 15 other errors, involving inconsistency of year of publication between text and references list, repetition of a reference listing or omission of the reference source were common errors. So many elementary mistakes are really unacceptable in a PhD thesis, reflecting poorly on the candidate's ability to concentrate on details.*

Where the examiner explores all of the following: the candidate's application of the literature and its role in the thesis, reference to theoretical contribution and the candidate's immersion in, and depth of understanding of the literature the text is coded at sub-category "Utilisation and/or theoretical application" (3 2 3).

> *One important aspect that a PhD should demonstrate is mastery of the relevant literature in both awareness and its potential application. X's reflection upon the meaning and direction of the procurement research, the authors cited inadequacies and limitations expressed in their work as well as in informal conference discussion all are well captured in this chapter. Placing the literature in X's chosen framework makes sense and adds value to the work of those he cites.*

Whereas the coverage of the literature was commented on in 53 per cent of the examiner reports, inaccuracies and utilisation and/or theoretical application were mentioned in 23 and 24 per cent respectively. 'Systematic acquisition and understanding of a substantial body of knowledge' is a generally accepted learning outcome of doctoral study (Shaw & Green 2002, p.117). In light of this, one might expect all examiners to comment on the literature, whether the review has been handled well or poorly. Clearly this is not the case. Nonetheless examiners do treat the literature as a critical indicator of the quality of the project reported in the thesis. This was demonstrated in the analysis of examiner reports on theses that were required to be revised and resubmitted (Lovat et al., 2002). From a comparison of reports where revision was and was not required, it is clear examiners expect candidates to have grasped the full extent and implications of the pertinent literature.

> *There does not appear to be any attempt to interpret the literature and provide a personal perspective on the information given. Rather it is a listing of known facts. Nor is there any attempt to provide a conceptual framework for the research to be undertaken. It is thus difficult to determine if the candidate has a comprehensive grasp of the subject under study, although the range and selection of Reference material is quite comprehensive.*

The absence of these qualities is significantly correlated with a recommendation for revision or outright failure. These findings add weight to those of Bruce (1994, 2001) that we need to

know a great deal more about how students approach and understand the literature review process as well as how they develop the skills to conduct an adequate review.

2.2 Approach (3 3)

Approach encompasses examiner comment on how the candidate approached the research, their methods and design, and variations in design or reflections on method. This category covers the nature of the experiments, instruments and mechanics of approaches to analysis (for example, the computer software employed) through to reasons given for their use. Further, this category also captures any comments on methodological and epistemological position taken in the thesis.

> *The thesis provides a thorough analysis of six case studies using quantitative and qualitative data and techniques. In discussion of the strengths and weaknesses of case study methodology is comprehensive (sic) and every effort is made to address the shortfall of this method of collecting and presenting a variety of data, both quantitative and qualitative.*

One of the assumptions the research team made in the early stages of the initial development of coding categories was that examiners would comment at length on research design and appropriateness. This was not the case. While it was mentioned in 60 per cent of reports, the mean proportion of each report devoted to approach was only 7 per cent. When considered in the context of the reports as a whole, this finding raised the possibility that examiners enter the thesis examination at a point somewhere beyond the proposal stage, that is, they comment primarily on elements they feel they can influence.

2.3 Subject matter and findings (3 4)

This category captures everything about the topic or subject matter and findings of the thesis under examination. Examiners very rarely described in detail what was in a thesis, over and above describing its scope. If examiners were not commenting on scope, method, literature or communication they were discussing the subject matter of the thesis primarily as results or findings. Such information is found in almost 90 per cent of reports. Such a strong focus by examiners on results is not surprising given that such subject matter is the reason for a thesis. On the rare occasion an examiner listed (without elaboration) what was in a chapter or the appendices, such comment was coded in a category labelled "ingredients" (3 10 3).

On average, the primary category (3 4) captures around half of total report text and it is to be expected that this material will yield important insights about disciplinary expectation. To undertake analyses by discipline will require access to discipline-based expertise and that expertise is built into a later stage of the study when we will call on disciplinary panels to explore examiner comment from a specialist perspective. In the core coding stage it was deemed practical to attempt no more than the accurate identification of the relevant text units devoted to discussion of subject matter. Nonetheless, it soon became evident that there were two levels worth identifying – one where the examiner concentrated on the text at hand (3 4 1), and another where the examiner added something new and possibly in detail (3 4 2).

The category Analysis and Reporting (3 4 1) captures all text covering the candidate's findings and the results of analysis and interpretation as well as the examiner's 'spin' on, or general response to, the candidate's endeavours. That is, the examiner conveys the way in which findings have been analysed and reported. Reporting of results incorporates the clarity and adequacy of the reporting and presentation of findings, including the examiner's discussion of the candidate's attention to validity, reliability, veracity, robustness, accuracy, strengths and weaknesses, and so on.

> *Throughout the exposition of the results the candidate also provides useful discussions on the limitations of the methods and instruments used to measure key variables. These discussions provide important information for other researchers in the field, particularly those engaged in research with disabled people. The discussions are detailed and well written and provide excellent information on key areas such as the usefulness of particular instruments and the effects of using mixed methods.*

The category 'topic related issues' (3 4 2) captures text where the examiner deals with specific issues in some detail, conveyed by their providing extended comments or reflection, and bringing their own knowledge and interests into play. These comments typically extend

beyond the specific subject matter of the thesis and introduce new information from the examiner's own corpus of knowledge, methodological skills and recent reading. Examples would be where an examiner cites the findings of a recent study, or highlights the way they would have gone about a particular step in the analysis, and possibly re-interprets the data. Lists of details of new references for the candidate also fall within this category.

As a case in point, consider the Alfven wave equation for a polytropic stratification a uniform vertical magnetic field (pages 54-59). The first column of Table 3.3 lists some of the discrete frequency eigenvalues for rigid boundaries applied at [Zb] and [Zt]. Bogdan & Cally (1997, Proc. R. Soc. Lond. A453, 943-961) solve the identical problem (in Section 3 of their paper) but allowing [Zt] and [Zb] to recede to the singular points of the ODE at zero and infinity.

2.4 Communicative competence (3 5)

The literature suggests that communicative competence (for example spelling, punctuation, mechanics, coherence and clarity) is a major topic in examiner reports on theses. Johnston (1997) found that 'almost without exception examiners commented on the writing and presentation' (p. 339). Hansford and Maxwell (1993), who investigated coursework masters theses, found the frequency of criticism of typographical and spelling errors to be noted by examiners in 42 per cent of theses, both writing style and format and presentation each to be noted in 22 per cent, grammatical and spelling errors in 17 per cent, referencing issues in 13 per cent, and bibliographic issues in 12 per cent of theses.

In the study reported here, 73 per cent of examiners mentioned substantial issues of communicative competence at least once, while 50 per cent identified specific corrections. Examiner comment that captures 'substantial issues' about 'communication and presentation' typically includes an holistic appraisal and/or summary of communication skills and is categorized as (3 5 1). These comments may not necessarily be critical. The examiner may state there is a major flaw, or many flaws, in communication or presentation, or conversely that there are none. They may convey frustration with weaknesses in communication or style throughout.

In contrast to Mr. X's impressive analytical work, its documentation, as well as the presentation of his results are marred by some editorial flaws. The structure of the thesis could have been tighter with less repetition and better integration of cause and effect. His literary style is patchy, ranging from rather eloquent prose to unedited laboratory jottings. There are some grammatical problems and too many cases of wrongly placed or missing punctuation marks.

Conversely examiners may praise the candidate for a well-developed conceptual map, well-honed argument, the absolute absence of errors and clear linkages throughout.

The thesis is very well written and presented; it is a pleasure to read. It has been carefully checked- at least, I could find no obvious errors. The candidate has dealt with quite complex concepts in a very clear and simple manner. The development and layout of the thesis are logical and easy to follow.

Category (3 5 2) captures editorial input. Many examiners will devote long sections of their report to editorial correction. On average in this study they devoted 5 per cent of text units or lines to substantial issues of communicative competence (351), but 8 per cent to editorial matters (typographical, basic mechanical, clerical or technical errors). A typical minor point is to say 'change this to that' (relating to a word, label, sentence, etc). Often the examiner will have a section of their report labelled typographical errors or editorial comment. Such sections contain 'fix-it' type comments and these are, by orientation, negative comments.

Examples*: it's should be its; delete one bracket; needs a space after "men"; Headings of chapters are Bold, in capitals, and underlined (all three), an ungainly style. It would look better if they were simply capitals.*

Sometimes the left margin justification is lost (as on p. 51), which gives a jagged appearance to the page.

This category does not capture minor errors in the bibliography or referencing. The latter are identified in (3 2 2).

To summarize, we can identify discrete themes and emphases in the text of examiner reports and these constitute the core categories for our analyses. The overwhelming emphasis is 'analysis and reporting' which is mentioned in 89 per cent of reports and, on average, constitutes about 43 per cent of a report, and in 44 per cent of cases the topic is taken up in some detail. Editorial comment appears in 50 per cent of reports but on average does not constitute a substantial proportion of the report (8 per cent). Substantial comment on communicative issues, while common (appearing at least once in 73 per cent of reports), accounts for even less of the report (5 per cent). The design and methods in and of themselves are mentioned in 60 per cent of reports, but account for only 7 per cent of total text on average, so depth of comment on this aspect is rare. The contribution of the thesis is mentioned in 70 per cent of cases, but accounts for about 9 per cent of total text, which is more than the text devoted to the literature. The ethics of the research are barely mentioned which, given the attention paid to the issue in Australian universities, is quite surprising. The lack of attention may well be related to the same reason that relatively little comment is devoted to research approach, namely the 'entry point' of the examiner. This is discussed further toward the end of the article.

Most of these comments about the different aspects of a thesis are also assessing or judging their merit, and it is to this layer of information that we now turn.

3. The Evaluative Elements in Examiner Comment

Examiner comment concerned wholly with inaccuracies in the literature (3 2 2) and editorial errors (3 5 2) is entirely negative in orientation. Most evaluative comment is not so straightforward, or specific to a particular feature of a thesis. The categories (5 1), (5 2), (5 3), and (5 4) reflect this and the underlying cultural complexity of doctoral examination. Some examiners may well see the thesis as the final product and provide an entirely summative assessment. Most see it as a transition point, and the thesis as a work-in-progress, that is, something that can be improved on. The primary category 'Evaluative elements' picks up on examiner feedback that is deeply embedded in the assessment of student knowing, situated within a unique professional discourse and personally idiosyncratic. Comment that falls under Category 5 can be positive or negative or structured in such a way as to encompass both. The meaning needs to be determined in the context of the entire report. The percentages of examiner reports that included at least one instance of each of the sub-categories under Evaluative Elements are shown in Table 2.

Table 2: Evaluative elements - occurrence of comment and proportion of total report (N=303)

Evaluative elements sub-categories	Occurrence in reports [a] %	Proportion of comment [b] %
Positive summative comment (5 1 1)	80	9.3
Neutral summative comment (5 1 2)	52	3.1
Negative summative comment (5 1 3)	16	1.0
Formative instruction (5 2)	68	28.6
Instructional commentary (5 3 1)	56	9.2
Prescription (5 3 2)	56	7.5
Other judgment (5 4)	88	14.3
Resubmission/plagiarism (5 10 2)	1	0

[a] The sub-category occurs at least once in a report.
[b] The proportion of comment from each sub-category as a percentage of the total report

There are four categories of evaluative comment that emerged from the reading and analysis of the reports. One type of judgement is summative (5 1). These comments tend to be stand-alone and succinctly sum up the examiner's feelings about the whole thesis or a particular sub-section of it. They pinpoint the qualities, or a quality, of the thesis, for example, 'this is a fine thesis', 'this is a well executed literature review', 'the analysis is deficient', and include general statements about the thesis as a whole. Such statements can be positive, negative or neutral. Positive summative comment pinpoints strengths. Negative summation pinpoints deficiencies and failings. However, there are often situations where the examiner endeavours to balance their comments. They may say something like '…on the one hand, [something is] good and on the other [that same something, or something else] is flawed. An examiner attempting to communicate a neutral position in a summary way may provide guarded praise or skirt around direct criticism by using an oblique turn of phrase.

> *Given the breadth of the task in hand one must say that the job is quite reasonably done. But still I wonder whether it would have been better to narrow the focus of attention, and cover the material in greater depth. I have some reservations about the thesis (quite strong ones in some areas), but my feeling is that the candidate has done just enough to warrant the award of the degree (with corrections).*

Summative comment (5 1) occurs in most reports, but constitutes a small proportion of the total report content, ranging on average from 9 per cent for positive summative comment (5 1 1) to 1 per cent for negative (5 1 3). Positive summative comment occurs in 80 per cent of reports, compared to neutral summative comment (5 1 2) which occurs in about half the reports, while 16 per cent of the reports contain negative summative comment. Occasionally an examiner will provide summative comment on thesis sections, major segments or even the thesis as a whole. It is rare for an examiner to restrict their evaluation to summative comment alone.

The next grouping of evaluative comment falls under the broad heading of *instruction*. That is, there is something in what the examiner is saying that is instructional. In the most basic sense the examiner is telling the candidate, and/or possibly the supervisor, to do something to improve the thesis or its products. It emerged that different qualities of instructional comment can be identified on the basis of how clear or unclear, and how complete or incomplete the total effect of the instructional text is. The most complete form of instructional comment or feedback is formative.

To be coded as Formative Instruction (5 2) the comment has to combine information for the reader (commentary), detail about how to proceed (direction and/or action) and involvement or engagement with the topic. This is why this comment is referred to as 'complete' and it is constructive, as in effect it comprises all the constituents for effective formative assessment. The thesis is, or elements of it are, treated as a work-in-progress that will benefit from constructive criticism or a different perspective on an issue or problem. There is a strong sense of there being an intended exchange of information in a teaching and learning sense. The examiner assists and challenges the candidate by providing new lines of information or possibilities, focussing expertise, knowledge and skills on a particular facet of the work. The suggestions often clearly anticipate 'dialogue'. The candidate may be asked what they think (and such questions extend beyond the simply rhetorical) or to evaluate a suggestion. The examiner may explore alternative scenarios, explanations or even strategies for reporting.

Consider the following statements by examiners which, although brief, give a sense of the complexity and multi-layered nature of formative instructional comment. In the first example it is clear the examiner is engaged by the issue, they state their dilemma, and offer information in which there is embedded direction for re-analysing the problem.

> *I have a philosophical argument…there is no nucleation energy barrier which implies that nucleation can occur easily. If that is the case, melt undercooling should actually be low. However all your measurements … show that melt undercooling is directly proportional to peak heat fluxes … Is there a critical cooling rate below which undercooling is actually lowered … What are your thoughts on this?*

In the next example there is more specific direction.

> *It is impossible to have X and Y vanish at both boundaries as stated in the middle of page 48 except for the trivial solution … As stated these boundary conditions over determine the system. I believe that the tabulated eigenvalues arise from somewhat different, although clearly admissable, boundary conditions. This subsection should be cleaned up to reflect what was actually computed*

There is other comment provided by examiners that indicates the thesis can or should be better but, in contrast to comment that is formative, this other instructive comment lacks the element of academic engagement or involvement. In this category 'Other Instruction' comment is either restricted to instructive commentary alone (that is, without extending to suggested or inferred action), or to instruction for action alone or with little or no commentary. It is on this basis the comment can be allocated into two further sub-categories.

'Instructive commentary' (5 3 1) is comment that conveys the sense of instruction. It is advice without adequate context or structure. It may be the kind of comment that supervisors and examination committees dismiss from their deliberations because the points have a slippery quality – a general vagueness, or it may not be possible to act on such advice. There can be a kernel of potentially useful knowledge that is not sufficiently extended to provide a clear understanding, or a statement of such breathtaking scope that it cannot be linked to anything other than the basic interpretation that the thesis or elements of it should have been different. Such comments may include rhetorical or broad questions, and be wide-ranging (suggesting or even showcasing the examiner's expertise). Four examples follow:

> *… the candidate did not integrate his findings into the existing literature, and did not provide much of his own analysis. Linking the findings into the literature situates them in their scholarly traditions, and demonstrates their wider applicability*
>
> *One drawback of the very broad scope of the thesis is that each of the topics discussed could only receive a relatively sketchy treatment. One could argue that each of the chapters could easily be a thesis topic in itself. However, it was obviously the author's choice to provide a broad overview …*
>
> *Why isn't the material of chapter 2 integrated with the material starting after page 140?*
>
> *The lack of representation, definition, decision rules , and analysis criteria for the model repeatedly constrains the ability of the candidate to apply structured argument to the theory-building discourse presented in the thesis.*

By way of contrast some examiners offer very detailed comment for action and little else. Their comments are essentially prescriptive. In contrast to instructive commentary (5 3 1) the category 'Prescription' (5 3 2) leaves the reader in little doubt about what the examiner expects them to do. The instructions are clear. 'Prescriptive' comment provides 'band-aid', 'fix it like this' type of directions without anticipation of challenge or negotiation. They are frequently comments that anticipate rapid closure on the thesis as opposed to more development. The examiner expects something to be added or enhanced but is not intellectually engaged with the content. Such comments may suggest the reader is working to a preferred formula or implied minimum standard.

The following are examples of prescriptive comment:

> *The opening argument needs to be re-argued. It needs operational definitions, critical comment on the two concepts of culture and management style. At least in this section if not throughout, updated references are required.*
>
> *It would have been pertinent to indicate that k was Faraday's constant … I would have referred to the degree of solar radiation rather than UV.*

The comments captured in the final category of judgement, 'Other Judgement' (5 4), typically convey a sense of responsiveness to the thesis by the examiner that range from the formal to the personal. 'Other judgement' comments are not summative, formative or instructive. Unlike summative evaluation which positions the thesis (in whole or part) along a continuum of 'very good' to 'very poor', the 'other judgement' category captures such dimensions as 'satisfaction' (or not), 'worth' (or not), 'effort' (or not) through to 'interest', 'displeasure', frustration, amazement, etc. In addition (5 4) comments lack the

weight/substance and/or succinct format of summative comment. This category of comments may very clearly exhibit the examiner's involvement or engagement with the thesis in their adjudication of its value or worth. The comment may be celebratory, passionate or declarative in nature, suggesting strong feelings about some feature or aspect of the thesis. They may also capture the reiteration of something positive or negative that the examiner has written at more length elsewhere in the report.

> *Again I congratulate the candidate on this excellent effort. This is among the most interesting theses I have ever read ...*
> *... several publications should emerge from this thesis ... it is surprising that there are none already.*
> *... an honest and at times courageous, attempt at a complex and difficult topic*
> *... a wonderfully accurate account of the Chinese philosophy of holism and demonstrates with many examples ... that Chinese medicine cannot be understood without reference to this holistic philosophy.*

Formative Instruction occurs in 68 per cent of reports, Instructive Commentary and Prescription each in 56 per cent, and 'other judgment' in 88 per cent. On average formative instruction accounts for 29 per cent of the examiner report while prescriptive comment is the shortest form of instructional comment (averaging 8 per cent of the total report).

Within the evaluative elements of examiner comment there are still finer distinctions in comment to be pursued. One facet already subjected to extended coding is the positive or negative orientation or tone of the comment. At the time of writing, such categories have been tested and what has been found is a preponderance of negative comment in 'Instruction-formative', a totally negative orientation in Prescription, and a mixture of positive and negative in the remaining two categories. If the editorial comment is added (which is negative as well), then a substantial proportion of what examiners say is negative in tone, despite a strong formative element in the assessment. To refer back to a point made much earlier in this article, one needs to ask why this is so, and what function does negativity perform?

4. Examiner Elements

Up to this point the paper has been concerned with features in the reports that are wholly about the thesis, but there is normally more to the examiner report than a focus on the particular thesis. There is also text that places the examiner at centre-stage. Such comments capture their approach to examination and report writing, the interpretation of their role and expectations, and their ways of conveying their response to the thesis to multiple readers. Some of this information falls strictly under the heading of personal reflection on, and understanding about, the processes involved in examination. The remainder reflects engagement with the thesis and the reading audience. The dominant dialogical features were captured from the outset by the coding process in order to juxtapose the manner and thrust of the comment against the themes and emphases highlighted by the examiner.

4.1 Examiner and process categories (2)

Under category (2), 'Examiner and Process', examiners may make comments about what they know or anticipate about examination and standards. They may comment on what they believe the candidate's institution expects, and they may talk about themselves. If such information is used to provide information about the examiner to justify their being chosen as an examiner or to justify the tone of their comments, it is coded under category (2).

The examiner may state quite specifically what they feel the candidate or process requires of them in terms of the type and depth of feedback. Such comments may cover knowledge of individuals including the candidate, self-set standards, or academic expectations pertinent to thesis examination. An examiner may comment about the limits of their expertise, or about how they have approached the task, possibly within the construction of the given guidelines. They may discuss the expectations or 'state of play' within the discipline or field, and its methods. They may acknowledge regional and institutional differences, and some of the text units may overlap with categories (3) and (4) if the points are elaborated, and argued. All such comments would seem to be intended to provide the context behind appraisal. The

following section describes and provides examples of text in category (2). The percentages of reports that include at least one mention of each of the sub-categories and the proportion of comment in each sub-category as a percentage of the total report are shown in Table 3.

Table 3
Examiner elements - occurrence of comment and proportion of total report (N=303)

Categories & sub-categories	Occurrence In reports [a] %	Proportion of comment [b] %
Examiner & process sub-categories (2)		
Providing personal & professional context (2 1)	59	6.6
Noting specific or anticipated criteria for examining (2 2)	71	6.8
Noting the 'model' or 'best type' thesis (2 3)	8	0.6
Mentioning supervisor or 'team' (2 4)	16	0.7
Dialogic elements sub-categories (4)		
Intellectual engagement (4 1)	29	3.5
Conversation (4 2)	69	8.9
Use of first person (4 3)	82	15.5

[a] The sub-category occurs at least once in a report.
[b] The proportion of comment from each sub-category as a percentage of the total report

In 59 per cent of reports the examiner provides 'personal and professional context' for the reader (2 1). On average there is about as much of this text in a report (7 per cent) as there is on significance and contribution (9 per cent).

The text captured under this sub-category provides personal grounding for the reader. For example, it may convey previous knowledge of the candidate or their institution, or the examiner's methodological strengths and knowledge of the field, or the developments within the field. This category could also indicate the examiner's perspective on the candidate's methodology or topic in relation to the above and would also capture text that compares the thesis with others they have examined, read or supervised. The category also captures reference to individuals or academics whom the candidate may or may not be expected to know, however, an exception here is the candidate's supervisor or team, because any such mentions are coded under 'Supervisor' (2 4).

> *The findings are further supported by the results coming out of my lab.*
> *... was discussing this with a colleague recently, and he was of the mind that this approach ...*
> *This is the essence of good science ... I was surprised to see that none of it had been published in peer reviewed publications ... after all, grant funding is contingent on a scientist's track record.*
> *I have recently become aware of the research of [X] in the area of ...*
> *I read an earlier version [of this thesis] and made some suggestions, these were effectively integrated into the final thesis.*

Where the examiner shows or claims technical or content expertise specifically in relation to the thesis and demonstrates it (engaging with the thesis), this is integral to assessment and will be identified in categories (3) and (5). There is much that examiners are expected to know as members of the academy that they bring to the assessment, including recent work, specific theories, etc. Much of this would fall within category (5). If such material is coded at (2), it is strictly because the information is offered to explain, frame or mediate the examiner's comments about certain aspects or the thesis or to 'position them' *vis a vis* the field.

Examiners convey what they expect from the thesis in 71 per cent of reports. This text is categorised as (2 2) the [application of] 'specific and anticipated criteria'. This sub-category also identifies reference to any specifications provided by the candidate's institution to guide the process, or reasons given why the examiner makes note of certain errors or inconsistencies, or why they instruct the candidate to make a change.

If the examiner goes further and comments on characteristics of a 'model PhD', this is coded at (2 3). Such comment occurs very rarely. In such comment the examiner refers to the expectations of a thesis held by the academy. Occasionally a report includes a brief lecture that sets the scene. Examples range from quotes from the literature about what is expected of a PhD, through to statements such as '… this is absolutely the top PhD I've read …', or '… outstanding; every aspect is attended to with thoroughness and expertise …'.

4.2 Dialogic elements categories (4)

Even the categories (2) (3) and (5) outlined above do not do justice to the richness of the text they capture. The primary category (4) was devised to tag dialogic elements in the text to facilitate extended analysis of disciplinary knowledge, examiner role and the academic community's understanding of what a PhD means.

It is anticipated that the examiner reports archived in this study will be an important source for researchers in assessment and higher education for many years, and will contribute an important, but usually inaccessible, dimension to the corpus of academic discourse.

In a situation such as a final examination, one might expect the examiner to evince distance. It is a mark of the extraordinary nature of PhD assessment that so little distance is evident in the examiner reports. It is fascinating to identify changes in the discourse where an examiner switches roles and direction in a way that anticipates a change in the audience. The conversational tone in 69 per cent of the reports (4 2) and the use of the first person (4 3) in 82 per cent of them was our earliest indication that many examiner reports contain something akin to parallel discourses. Given that accounts are highly context specific, the construction of academic identity is more in evidence and accessible in examiner reports than we would have anticipated in that medium.

When the reports were coded, the coders were not aware of the examiner ratings that were given with the reports. In 29 per cent of reports examiners seemed to show intellectual engagement with the subject matter being evaluated, and this led us to assume such engagement would reflect the top rating theses. Later we were to find out it was not the case. 'Intellectual engagement' only occurs in cases where the examiner explores a topic at some length, arguing with the ideas, extending the analysis, musing to self, and talking to the imagined readers about connected aspects or alternative interpretations. The latter has all the hallmarks of an imaginary dialogue or conversation, which is also captured in category (4 2), 'Conversation'.

Where conversation occurs in a report, it may exhibit overt or subtle changes as the examiner moves from one role to another (arbiter, colleague, supervisor), and one audience to another. Candidates are assisted and often chided, supervisors may be admonished, administrators addressed, and colleagues hailed. The form of address can move from the formal to the informal, from aloof distance to familiarity.

> *In any event, if you are going to cite this line from Alexander, which makes clear reference to the Stoic definition of the possible, it seems to me requisite that you discuss in much more detail their treatment of modality. In general, it seems to me that you are too quickly assimilating a couple of quite different problems.*
>
> *The reference to performance-based assessment is ridiculously short. Much has changed since then and the thesis would be much better if the more recent literature were included. My assessment reflects tolerance of the fact that doing this was probably not mandated, but I think that the candidate and supervisors should be aware that this defect makes the result a very weak thesis for 1999-2000.*
>
> *I find myself in agreement with the majority of the candidate's analyses and conclusions. However, in a number of instances I was not so convinced. I do not wish to list and*

specify my agreements and disagreements here one by one, because I do not believe an evaluation of a thesis should depend on the degree to which the reader(examiner) is convinced by it, provided the conclusions are based on a sound scientific approach (methodology) and critical argument.

In her study of examiner reports, Johnston (1997) provided a range of illustrations to show that the report was more than a treatment of the intellectual endeavour exhibited in a thesis. She highlighted the small amount of direct reference to the supervisor, but touched on the implicit references to supervisory responsibility and quality, and explored these references as 'messages'. Examiner feelings and application of standards are also raised in her findings. These layers have been identified and captured in our text analysis. We also added another element that Johnston canvassed, namely report structure, but in a way that can be utilised to identify and evaluate the communicative qualities of the examiner reports.

5. Organization Elements: Report Structure

The categories used here are structural ones that primarily capture how the text is grouped under examiner-designated headings. It might be expected that heading type would indicate what examiners believe is important, however, no such relationship was identified in this study. Indeed, there is no standard structure or style evident at all. Examiners used a variety of fonts, word processing features and page settings, ranging from standard to unconventional. Such features appear to reflect skills in word processing, editing, formatting, and time pressures. For the institution studied, examiner guidelines were vague and the guidelines that were given were not apparent in the structural features of the reports. Structure was individualistic, showing considerable variety in sequence and sectioning particularly evident in use of sub-headings. There are seven main section types identified in the reports, often reflected in sub-heading titles and coded under primary level category (1). The most typical sub-heading is the 'general' type, e.g., 'general comments', 'minor comments', 'main points', 'final points', 'summary', and most offer very little direct guidance to the reader. Others provide a mixture of 'general' and 'error' sections, and a number of reports dealt with points chapter by chapter. Very few of the reports were organised by specific themes, topics or methods. On the whole the standard of presentation in the examiners reports would in the majority of cases fall short of what emerged from the reports as the standard expected from candidates.

6. Discussion

The content and conceptual analysis that characterise the core coding phase have allowed us to draw a number of conclusions, some more tentative than others and awaiting further testing through cross-institutional comparison, and cross-disciplinary approaches to the text.

6.1 Comment type and attributes

While examiner reports are very individual documents they generally cover common ground. The main areas or aspects of the thesis that attract comment are the scope, significance and contribution of the thesis, including publications, the approach or method taken, treatment of, and engagement with, the literature or scholarly reading, the analysis undertaken, and the communicative competence demonstrated by the candidate (including basic mechanical, typographical and referencing errors). The breadth of these areas negates the possibility of a predominance of one approach to analysis and reporting. It is also an indication of the marked differences in the reports that while 89 per cent of examiners comment on analysis and reporting, 11 per cent do not, and that 31 per cent do not mention significance and 47 per cent do not comment on the coverage of the literature in the field; 50 per cent provide no specific editorial comment, although the majority refer to substantive issues of communication.

Although some examiners spend a section of the report describing the thesis and provide evaluative comments separately, it is more typical that remarks on the content of the thesis overlap with advice about how to improve the document or publications that emanate from it. In fact the substance of the findings engage the examiner so extensively and intensively, that

the typical examiner addresses the candidate directly, engaging them in some dialogue about their findings.

Another finding of note relates to the qualities of instruction in the reports. Examiners present their judgments in various forms. There are unambiguous summative statements and short, sometimes pithy, general statements, which indicate their position(s) on the qualities of the thesis. However, most of the evaluative comment is coupled with some variant of instruction. A substantial proportion of examiner comment is formative and presumes intellectual engagement by the reader with the examiner's concerns, queries, ideas, and knowledge. Another mode of instruction is prescriptive. In this mode the pedagogy is rule-based and examiners anticipate agreement and compliance, not critical engagement. In such a situation the examiner may provide the answers to their own queries, in whole or in part – prescribing the direction for the candidate or supervisor with 'you must do', 'should do', or 'need to' statements. These statements are in addition to editorial comments. Editorial comment in the examiner report seeks the correction of mechanical, typographical or referencing errors and the examiner usually supplies the verbatim correction. There is one further form of instruction, but this takes the form of a vague, loose or general commentary with tenuous links to action. Hence instructive commentary is not necessarily helpful or constructive.

In summary, we find that examiners engage with a range of thesis features, but mostly with the analysis and interpretation of the findings, and then mostly in the spirit of instruction. What is not clear from the text analysis alone is the weighting given to instruction in the final assessment or rating of the thesis. Moreover, an intriguing question posed by the findings is where does the examination begin? Is there an entry point indicated by the relative emphases in examiner comment?

6.2 Examiner emphasis

In terms of proportion of text devoted to the various areas of the thesis, significance, and project method and design receive scant attention by examiners. Significance may be based on the quality of the thesis and/or project or its specific contribution (Noble, 1994, p. 101). The coding identified both these aspects in the one category. That 69 per cent said something about significance, and 60 per cent said something about approach and methodology indicates the high profile of these features in the doctoral examination process, but the types of comment undermine their importance. Mostly they were comments of the surface or descriptive type. They did not 'engage' the examiner. Moreover the type of evaluative comment allied to these areas was mostly summative and non-instructive. This is in stark contrast to the often lengthy and minutely detailed engagement with the findings and analysis.

Design and significance elements are more commonly associated with earlier stages of the project and defence of the proposal. Examiners were far more active in those areas where their comments could make a difference, that is, where change and improvement was an option. This in turn suggests their main entry point into the process appears to be defined by the possibility of involvement. Instruction is intended to raise the thesis or publications from the thesis to an 'appropriate' level, defined primarily by what has not been done. Some tentative conclusions can be offered about this tendency in relation to how examiners utilize the report.

6.3 Use of the report – audience and role

Most students want to know what examiners are looking for, but what examiners look for and what they privilege in their reports may be different things. The difference is significant. Examiners may not feel it is appropriate to demand certain changes, or their silence on certain issues may suggest there is nothing of substance to say (even in a positive sense). What they do and do not say is very important in a number of ways. On the one hand their comments may be used by a committee making the final decision on a thesis in an attempt to resolve differences in examiner rating, or weighed to resolve differences in examiner emphasis. The comments will also be judged for their credibility and usefulness and may, as a consequence, receive either an enhanced or diminished weight in the committee's decision on the thesis. Given the complex set of expectations in play, it is little wonder that, contrary to a common

sense understanding of 'final examination', thesis examination is not unequivocally 'summative', and has multiple purposes, some of which have nothing to do with assessment.

Examiners use the report on the PhD in at least five ways: to judge, instruct, amend, mentor, and confer or confirm membership of the academy (perhaps both their own and the candidate's membership). These uses are allied to certain roles. The roles could be described as assessor-arbiter, mentor-colleague and supervisor-instructor.

> *Mentor-colleague:* This role is characterised by involvement, peer-type acceptance, a concern to assist through providing critical support aimed toward encouraging independence and personal growth of the candidate.
>
> *Supervisor* : This role is more the teaching role that involves the individual taking responsibility for bringing the individual to an acceptable standard. There is a sense of having control over outcome and taking the opportunity to exercise that control.
>
> *Assessor-arbiter*: This is the more traditional examiner role, distanced and deliberative, distinguishing between correct and incorrect, accurate and inaccurate and applying specifiable criteria. It is a role that accepts the existence of in-built restraints in the assessment situation (that is that certain decisions or judgements cannot be undone) and that the outcome cannot be changed.

The identification of these roles within the assessment process has profound implications for supervisor and examiner training.

The high proportion of report text devoted to formative instruction – to instruction and amendment of the thesis - suggests the instructor or supervisor role is more dominant than the arbiter role as reflected in the use of the report. Does this suggest that candidates are generally seen as colleagues, as ex officio members? Is the sting of mostly negative and critical comment the equivalent of initiation? Examiners are usually supervisors themselves and appreciate the effort involved in a PhD thesis, whatever the standard reached by the candidate. Jackson and Tinkler (2001) and Mullins and Kiley (2002) noted that examiners described their role as gatekeepers, with an obligation to maintain standards. Yet, the desire to set the bar high is tempered by the desire to encourage and instruct (mentor and advise). There are summative comments that clearly set forth a judgment, but formative comments that anticipate that minimum standards will be reached and higher ones aspired to.

The instructor role has its own peculiar alchemy. For example, when examiners are at their most prescriptive they are at their least flexible and most judicial. Nonetheless examiner prescription is harshest at the surface. Fix-it type prescriptive comment typically lowers the bar or at least bends it and as such signifies softening or abandonment of the arbiter role. Moreover the target readership for prescriptive comment would seem to be as much the supervisor as the candidate – either in admonition or commiseration. With prescriptive comment the examiner also edges into the supervisory role, and with intellectual engagement into a mentoring or collegial role. The mutable nature of examiner role presents quite a challenge to the identification of consistent indicators of thesis quality in the reports. It has also provided an invaluable perspective on comment frequency and emphasis.

In this phase of the text analysis, key areas of doctoral examiner comment, their frequency and relative emphasis have been identified. The coding process illuminated examiner role and examiner entry point and these conceptual categories were subsequently fleshed out by the interrogation and exploration of recurring patterns in the data. Other researchers have noted the overlap between the experiences of supervision and examination (Mullins & Kiley, 2002) and the prevalence of formative instruction, but our findings did not support earlier findings by Johnston (1997) that examiners were substantially engaged in editorial activity, or that communicative competence was a significant problem (Noble, 1994).

Why examiners engage with the thesis as they do, and how their comment predicts their final decision are lines of inquiry that will be pursued in future analyses.

REFERENCES

Advisory Board for the Research Councils (1992). The Nature of the PhD. A Discussion Document, Office of Science and Technology. U.K.

Bruce, C.S. (1994). Research students' early experiences of the dissertation literature review, *Studies in Higher Education, 19* (2), 217-229.

Bruce C. (2001). Interpreting the scope of their literature reviews: significant differences in research students' concerns, New Library world 102 (1163/1164) (pp.158-165), http://www.emerald-library.com/ft

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., Lovat, T. & Monfries, M. (2001). Examiners as Teachers: the instructional content of PhD reports. BERA Conference, Leeds UK, 13-15 September.

Jackson, C. & Tinkler, P (2001). Back to basics: a consideration of the purposes of the PhD viva, *Assessment & Evaluation in Higher Education, 26* (4), 355-366.

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

Kouptsov, O. (1994). *The Doctorate in the Europe Region*. CEPES UNESCO.

Noble, K. A. (1994), *Changing Doctoral Degrees: An International Perspective.* Buckingham: The Society for Research into Higher Education & Open University Press.

Lovat, T., Holbrook, A., Bourke, S., Dally, K. & Hazel, G. (2002). Examiner comment on theses that have been revised and resubmitted. Paper presented at the AARE Conference, Brisbane, 1-5 December. http://www.aare.edu.au/index.html/LOV02282

Mullins, G. & Kiley, M. (2002). 'It's a PhD, not a Nobel Prize': how experienced examiners assess research theses. *Studies in Higher Education, 27* (4), 369-386.

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 Diversion.* (pp. 129-136). Melbourne: HERDSA.

Richards, T. L. & Richards, L. (1994). Using Computers in Qualitative Research. In N. K. Denzin & Y. S. Lincoln (Eds). *Handbook of Qualitative Research*. (pp. 445 - 462).Thousand Oaks Calif: Sage.

Shaw, M. & Green, D. (2002). Benchmarking the PhD – a tentative beginning. *Quality Assurance in Education, 10* (2), 116-124.

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.

In addition to the core academic team, the authors would like to acknowledge the assistance of the following individuals in the coding trial process: Mr Gavin Hazel and Dr Melissa Monfries, University of Newcastle, Australia.

APPENDIX

1 REPORT ORGANISATION
- 1 1 GENERAL SECTION
  - 1 1 1 first
  - 1 1 2 last
  - 1 1 3 middle
- 1 2 CHAPTER SECTION
- 1 3 ERROR SECTION
- 1 4 REQUIRED AMENDMENTS
- 1 5 CONCEPT,TOPIC
- 1 6 METHOD, FINDINGS
- 1 7 DIRECTIONS FOR READER
- 1 10 OTHER 1

2 EXAMINER AND PROCESS
- 2 1 PERSONAL & PROFESSIONAL CONTEXT
- 2 2 SPECIFIC & ANTICIPATED CRITERIA
- 2 3 THE MODEL PhD
- 2 4 SUPERVISOR
- 2 5 RECOMMENDATION
- 2 10 OTHER 2

3 ASSESSABLE AREAS COVERED
- 3 1 SCOPE, SIGNIFICANCE & CONTRIBUTION
  - 3 1 1 scope
  - 3 1 2 significance & contribution
  - 3 1 3 publications arising
  - 3 1 4 existing publications
- 3 2 REVIEW OF THE LITERATURE
  - 3 2 1 coverage
  - 3 2 2 inaccuracy
  - 3 2 3 utilisation/ theoretical application
- 3 3 APPROACH
- 3 4 SUBJECT MATTER, FINDINGS
  - 3 4 1 analysis and reporting
  - 3 4 2 topic related issues
- 3 5 COMMUNICATIVE COMPETENCE
  - 3 5 1 substantial issues
  - 3 5 2 editorial issues
- 3 10 OTHER 3
  - 3 10 1 ethics
  - 3 10 2 other
  - 3 10 3 ingredients

4 DIALOGIC ELEMENTS
- 4 1 INTELLECTUAL ENGAGEMENT
- 4 2 CONVERSATION
- 4 3 FIRST PERSON
- 4 10 OTHER 4

5 EVALUATIVE ELEMENTS
- 5 1 SUMMATIVE
  - 5 1 1 positive
  - 5 1 2 neutral
  - 5 1 3 negative
- 5 2 FORMATIVE INSTRUCTION
- 5 3 OTHER INSTRUCTION
  - 5 3 1 commentary
  - 5 3 2 prescriptive
- 5 4 OTHER JUDGEMENT
  - 5 4 1 positive
  - 5 4 2 neither
  - 5 4 3 negative
- 5 10 OTHER 5
  - 5 10 1 other
  - 5 10 2 resubmission and/or plagiarism