



E I P

**T**he 1999 Postgraduate  
Research Experience  
Questionnaire

---

01/4

John Ainley

Australian Council for  
Educational Research

---

Evaluations and Investigations Programme  
Higher Education Division



Department of Education,  
Training and Youth Affairs





Department of Education,  
Training and Youth Affairs

# The 1999 Postgraduate Research Experience Questionnaire

John Ainley

Australian Council for Educational Research

01/4 April 2001

Evaluations and Investigations Programme  
Higher Education Division

---

## Instructions for pdf navigation

- Use the arrows on the Acrobat menu bar to navigate forwards or backwards page by page
- Alternatively, use the arrow icons on your keyboard to navigate through the document.
- To enlarge the viewing screen either:
  - use the magnifying glass by clicking on the area you wish to enlarge or by forming a marquee over the area you wish to view (ie. hold the mouse button down and drag the magnifying glass over the area); or
  - use the view options menu bar at the bottom of the Acrobat screen.
- To pan out from the page, hold down the option button on your keyboard to change the +ve symbol on the magnifying glass to a -ve symbol , then click the mouse.
- To search for a word or phrase use the binoculars icon on the menu bar.
- The Contents pages are live, ie. if you click on a topic you will go to that page.
- You can return to the Contents page by clicking your mouse on 'Contents' on the top of each page.

© Commonwealth of Australia 2000

ISBN 0 642 45724 7

ISBN 0 642 45725 5 (Online version)

DETYA No. 6671 HERC01A

This work is copyright. Apart from any use as permitted under the *Copyright Act 1968*, no part may be reproduced by any process without permission from Ausinfo. Requests and inquiries concerning reproduction and rights should be addressed to the Manager, Legislative Services, Ausinfo, GPO Box 84, Canberra ACT 2601.

The report is funded under the Evaluations and Investigations Programme of the Department of Education, Training and Youth Affairs.

The views expressed in this report do not necessarily reflect the views of the Department of Education, Training and Youth Affairs.

# Contents

Contents .....	.iii
Acknowledgments .....	.vii
1. Introduction .....	.1
1.1 Background .....	.1
1.2 The Postgraduate Research Experience Questionnaire .....	.2
1.2.1 Development .....	.2
1.2.2 Format .....	.3
1.2.3 The scales .....	.3
1.2.4 Within-programme variability .....	.4
1.3 The 1999 survey .....	.4
1.3.1 Response rates .....	.5
1.3.2 Issues arising from the numbers of respondents .....	.6
1.4 About this report .....	.6
2. National results from the PREQ .....	.9
2.1 Responses to items .....	.9
2.1.1 Distributions of responses .....	.9
2.1.2 Summary statistics .....	.12
2.2 Groups of items and scales .....	.16
2.2.1 Summary statistics for groups of items or scales .....	.17
2.2.2 Scale summary statistics for the PREQ .....	.17
2.3 Structure .....	.18
2.3.1 Exploratory principal components analysis .....	.18
2.3.2 Confirmatory factor analysis .....	.20
2.3.3 Scale reliability .....	.21
2.3.4 Correlations between scales .....	.22
2.4 Summary .....	.23

3.	Background characteristics of PREQ respondents in 1999 . . . . .	25
3.1	Gender . . . . .	25
3.2	Age of respondents . . . . .	26
3.3	Language background and resident status . . . . .	28
3.4	Mode of study . . . . .	30
3.5	Summary . . . . .	31
4.	Differences between levels and field of study . . . . .	33
4.1	Level of study . . . . .	33
4.2	Broad field of study . . . . .	35
4.2.1	Supervision . . . . .	37
4.2.2	Skill development . . . . .	38
4.2.3	Intellectual climate . . . . .	38
4.2.4	Infrastructure . . . . .	39
4.2.5	Thesis examination . . . . .	40
4.2.6	Overall satisfaction . . . . .	41
4.3	Programme level and field of study . . . . .	41
4.4	Summary . . . . .	46
5.	Institutional patterns . . . . .	47
5.1	Overview . . . . .	48
5.2	Procedures . . . . .	49
5.2.1	Allowing for other influences by estimating residual scores . . . . .	49
5.2.2	Generating institutional measures . . . . .	50
5.2.3	Comparing institutional means . . . . .	50
5.3	Results . . . . .	51
5.4	Summary . . . . .	59
6.	Conclusion . . . . .	61
	References . . . . .	63

Appendix A The Postgraduate Research Experience Questionnaire 1999 . . . . .65

Appendix B Issues in the analysis of institutional means . . . . .67

**Tables and figures**

Tables

Table 1: Characteristics of respondents to PREQ 1999 survey and postgraduate completions in 1998 . . . . .7

Table 2: Distribution of responses and descriptive statistics for PREQ 1999 Items: respondents from Australian universities . . . . .11

Table 3: Percentage agreement with items on the PREQ 1999: Respondents from Australian Universities . . . . .14

Table 4: Summary statistics for PREQ sub-scales for all respondents . . . . .18

Table 5: Exploratory principal components analysis of the 1999 PREQ items . . .19

Table 6: Fit indexes for multi-factor confirmatory factor analysis of the PREQ . . .20

Table 7: Coefficients from confirmatory factor analysis of the 1999 PREQ items .21

Table 8: Reliability estimates of 1999 PREQ scales . . . . .22

Table 9: Correlation coefficients between PREQ scale scores . . . . .23

Table 10: Mean percentage agreement with PREQ sub-scales for male and female respondents . . . . .26

Table 11: Mean percentage agreement with PREQ sub-scales by respondent age groups . . . . .27

Table 12: Mean percentage agreement with PREQ sub-scales for non-English speaking background and English speaking background respondents . . .29

Table 13: Mean percentage agreement with PREQ sub-scales for Australian and overseas resident respondents . . . . .30

Table 14: Mean percentage agreement to PREQ sub-scales for full-time and part-time respondents . . . . .31

Table 15: Mean percentage agreement with PREQ sub-scales for research masters and doctor of philosophy respondents . . . . .34

Table 16: Mean percentage agreement with PREQ sub-scales by broad field of study . . . . .35

Table 17: Percentage respondents in different programmes by field of study . . . . .42

Table 18: Percentage agreement scores by level of programme and field of study . . . . .42

Figures

Figure 1: Institutional response rates to the PREQ in 1999 . . . . .5

Figure 2: Percentage agreement for PREQ items . . . . .15

Figure 3: Mean percentage agreement for supervision by field of study . . . . .37

Figure 4: Mean percentage agreement for skill development by field of study . . .38

Figure 5: Mean percentage agreement scores for intellectual climate  
by field of study . . . . .39

Figure 6: Mean percentage agreement for infrastructure by field of study . . . . .40

Figure 7: Mean percentage agreement for thesis examination by field of study . .40

Figure 8: Mean percentage agreement for overall satisfaction by field of study . .41

Figure 9: Profiles of PREQ scales by programme level and field of study . . . . .45

Figure 10: PREQ supervision scores for institutions . . . . .52

Figure 11: PREQ skill development scores for institutions . . . . .53

Figure 12: PREQ intellectual climate scores for institutions . . . . .54

Figure 13: PREQ infrastructure scores for institutions . . . . .55

Figure 14: PREQ thesis examination scores for institutions . . . . .56

Figure 15: PREQ goals and expectations scores for institutions . . . . .57

Figure 16: PREQ overall satisfaction scores for institutions . . . . .58

# Acknowledgments

The 1999 Postgraduate Research Experience survey was commissioned by the Graduate Careers Council of Australia.

---

## From ACER

The study was undertaken with the assistance of a project Advisory Group and staff from the Higher Education Division of the Commonwealth Department of Education, Training and Youth Affairs; their advice and feedback is gratefully acknowledged.

ACER also wishes to acknowledge the efforts of the Project Director, Roger Bartley, the Survey Project Director, Bruce Guthrie, and staff from the Graduate Careers Council of Australia.

We especially thank Dr Gerald Elsworth from the Faculty of Education at the University of Melbourne for his ideas and advice on the analyses used in the investigation. Dr Elsworth conducted the confirmatory factor analyses reported in Chapter 2. Ms Kylie Hillman contributed to a number of the analyses contained in the report.

---

## From the GCCA

The Graduate Careers Council of Australia wish to thank the Department of Education, Training and Youth Affairs for funding the *1999 Postgraduate Research Experience Questionnaire* through their Evaluations and Investigations Programme.

The Graduate Careers Council of Australia coordinated the acquisition of PREQ data from Australian and New Zealand universities. We wish to acknowledge the role of these participating universities—in particular the Deans of Faculties, research directors, survey managers and careers service staff who provided valuable support to the project. To the postgraduate students who completed survey forms we express our sincere appreciation.

# 1. Introduction

The Postgraduate Research Experience Questionnaire is designed to provide information on the educational experiences of students in higher research degree courses in Australian institutions.

---

## 1.1 Background

In Australia, as in other countries, increased attention on the quality of provision in higher education corresponded with an expansion of participation in higher education. In a situation of high demand for places in higher education, and limited resources with which to provide those places, the utilisation of resources becomes important. During the late 1980s a series of reviews of teaching in various disciplines (in Accounting, Engineering) and government enquiries argued the case for more systematic data gathering about teaching and learning in higher education. Government statements at this time crystallised the view that there was need for greater public accountability in higher education. It was argued that universities needed to be able to make judgements about whether the courses they provide are well taught, whether changes that are made to those courses impact on graduates' perceptions and whether graduates consider their experience of the courses to be rewarding and satisfying. In this discussion it was suggested that inferences about the quality of courses be facilitated by relative measures, comparisons between universities and faculties, schools and departments and comparisons over time. The *Performance Indicators Review Group* (PIRG) recommended that an instrument structured initially along the lines of the Course Experience Questionnaire be incorporated in, or administered in conjunction with, the GCCA graduate destination survey.

Since 1992, the Graduate Careers Council of Australia (GCCA) has included a Course Experience Questionnaire (CEQ) with its annual survey of graduate destinations: the Graduate Destination Survey. The CEQ is based on a theory that sees students' perceptions of curriculum, instruction and assessment as key determinants of their approach to learning and the quality of the outcomes of that learning. Through this questionnaire graduates record their opinions about a number of aspects of teaching and learning in the courses that they have completed. This instrument has captured graduates' perceptions of their higher education experiences, and is generally considered to be both robust and reliable. Twenty four items form five scales: Good Teaching, Clear Goals and Standards, Appropriate Workload, Appropriate

---

Assessment and Generic Skills. In addition it contains a single item concerned with Overall Satisfaction.

For a number of years, however, it has been recognised that the CEQ, which was originally developed for bachelor degree graduates, is inappropriate for the growing number of postgraduate research students in Australia. For this reason, in 1996 the Department of Employment, Education, Training and Youth Affairs asked the GCCA to develop a brief instrument in similar form to the CEQ which would gather data concerning the experiences of research higher degree graduates.

---

## 1.2 The Postgraduate Research Experience Questionnaire

The current (1999) version of the *Postgraduate Research Experience Questionnaire* (PREQ) was developed by the Australian Council for Educational Research (ACER) and the GCCA. It has a number of potential uses: providing a national overview of broad issues; setting a context against which institutions might undertake their own analyses and identifying examples of good practice (ACER, 1999).

### 1.2.1 Development

The GCCA was responsible for the initial conceptualisation of the instrument and the development of the individual items. ACER conducted analyses of the data from the first and second trial versions of the PREQ, leading to the selection of the 28 items that form the six sub-scales of the present PREQ. Twenty-eight universities, involving some 2336 postgraduate research students, participated in the first trial. In the second trial 29 Australian (and three New Zealand) universities participated, involving 1832 postgraduate research students. In both trials two forms of the instrument were tested, one using an agree—disagree scale and the other a satisfied—dissatisfied scale. Data analysis was carried out using exploratory principal components analysis, item response theory and multi-factor Structural Equation Modelling, where appropriate. The ‘agree’ version of the instrument proved far more robust. There were far fewer problems with missing data, and interpretation of factors was far more straightforward. Item response theory analyses successfully identified relatively homogeneous clusters of items on the ‘agree’ questionnaire that were clearly related to the underlying dimensions of the student experience. Multi-factor confirmatory factor analysis confirmed that the measures based on the six clusters of items did indeed identify separate constructs as described in the exploratory factor analysis. Readers are referred

to the report on the evaluation and validation of the trial PREQ (ACER, 1999) for a full discussion of the development of the instrument and sub-scales. More extended analysis of the PREQ trial data has been reported by Elsworth (1999).

### 1.2.2 Format

The PREQ consists of 28 statements about research students' higher degree experiences. Respondents indicate the extent of their agreement with each statement by selecting a number from one through to five (one standing for '*Strongly Disagree*' and five for '*Strongly Agree*'). Those who believe that the statement does not apply to them or their experiences may indicate so by selecting '*Does not Apply*'.

### 1.2.3 The scales

The purpose of the PREQ is to gather data concerning the experience of research degree (Masters by Research and PhD) graduates with regard to broad aspects of their studies. As with the CEQ, which uses 25 items, there is no attempt to be all-inclusive in terms of the experiences of the target group or in the range of possible experiences. There might be other dimensions that reflect particular circumstances of a field of study, a programme or an institution. Even within the areas addressed by this questionnaire there might be further issues to be probed in more specific studies.

The strength of the present approach is that it focuses on dimensions central to the postgraduate experience in most fields of study. The items included in the PREQ and the scales that they form were initially based on a review of the literature and discussions within focus groups of research higher degree students. They were revised on the basis of item performance during the first and second trials (see ACER, 1999 for greater detail of the scale development).

The scales that form the current version of the PREQ focus on the following six areas of research higher degree experience:

- Supervision;
- Intellectual climate;
- Skills development;
- Infrastructure;
- Thesis examination process; and
- Clarity of goals and expectations.

In addition there is, as in the CEQ, an Overall Satisfaction item. Appendix A contains a copy of the questionnaire. Chapter 2 provides information about the PREQ scales and its structure based on the current data.

#### 1.2.4 Within-programme variability

The emphasis of the PREQ is on the overall experience as a postgraduate research student. It is sometimes argued that the experience of postgraduate study is different for each individual, and that, dependent on each supervisor, there may not be systematic patterns within units (departments, schools or faculties) in institutions.

The argument is an extension of a general issue about studies of learning environments experienced in bachelor degree programmes (eg using the CEQ) or in schools. It is recognised that for most measures of learning environments there is variation among respondents within courses and institutions. This variation arises both because individuals may have different perceptions of the same situation and because there can be actual variations in the situation that is experienced within a field of study within an institution. Despite this, in university bachelor degree programmes, and in schools, it has been possible to establish shared or common perceptions of programmes and to relate those perceptions to characteristics of the course (Wilson, Lizzio & Ramsden, 1997).

Although it can be argued that the postgraduate research student experience might differ between supervisors even within the same department, there are other aspects to consider. Postgraduate research students may learn from each other, and from academic staff other than their own supervisor, depending upon the environment of the department in which they are enrolled. Indeed, the way supervisors interact with postgraduate students may be shaped by formal requirements of the department and by the less formal patterns of work across the department. In other words, the extent of within-programme variability is a matter of interest but it should not be assumed that the experience of postgraduate study is entirely supervisor dependent. Since the PREQ encompasses a number of dimensions it should be possible to explore whether within-programme variability is greater on some dimensions than others.

---

### 1.3 The 1999 survey

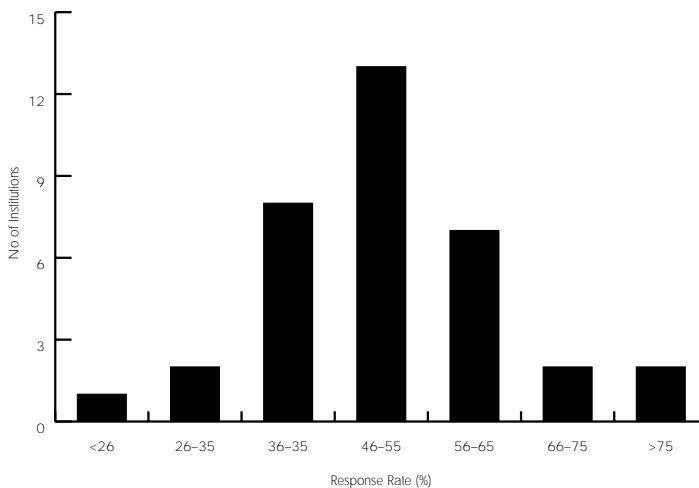
The PREQ was sent to approximately 4600 research higher degree graduates, who had completed either a Masters by Research or a PhD in the financial year 1998 to 1999. Responses were received from 2253 postgraduates in 35

Australian universities and the University of Waikato in New Zealand. There were 2225 respondents from Australian universities. The largest number of responses from an institution was 262 and the smallest was four. On average there were 63 respondents per institution.

### 1.3.1 Response rates

The overall response rate was 49 per cent. The average response rate for participating Australian institutions was just less than 50 per cent. Three quarters of the institutions had a response rate of 40 per cent or more and one quarter of the institutions had a response rate of 57 per cent or more. The distribution of response rates is shown in Figure 1.

Figure 1: Institutional response rates to the PREQ in 1999



The overall response rate is not above the minimum 50 per cent required by the AVCC guidelines for reporting of the CEQ results, but is similar to the response rates obtained by ACER (1999) in the developmental studies of the PREQ. It may be that there are some characteristics particular to this population of graduates that is limiting response rates to the PREQ. Findings from the PREQ 1999, therefore, should be interpreted with caution, due to the less than ideal response rate, as findings reported here may not be representative of the feelings of the entire postgraduate survey population.

The issue of potential bias from non-response affects how one deals with the accuracy of the data. The achieved set of responses is not a structured sample from the population and the estimation of errors based on the presumption that it is a sample is not appropriate. If the response rate were 100 per cent, sampling error would not be relevant, although one would still need to allow

for measurement error. The question is the extent to which one should allow for the potential bias introduced by non-response.

From the response data it is evident that there is little difference in response rate by level of study. For the 3141 PhD graduates who were sent a PREQ form, 1570 were returned, giving a response rate of 50 per cent for PhD graduates, whilst the corresponding response rate for Masters by Research graduates was 45 per cent.

As a check on the extent of bias in response, it is possible to compare the distribution of key characteristics among respondents to the PREQ survey in 1999 with postgraduate research students in general in that year. These characteristics were level of study, broad field of study, gender and age. Information relevant to that comparison is contained in Table 1. From those data it can be seen that the distribution of those characteristics among respondents to the PREQ matched the corresponding distribution among postgraduate research students. There was a slight under representation of engineering graduates. There was also an under representation of graduates aged 25 and below but that could reflect a real difference between the graduate population and the student population rather than any bias in response. Overall these data suggest that the respondents to the PREQ in 1999 did not reflect bias in these background characteristics.

### 1.3.2 Issues arising from the numbers of respondents

There are other issues that arise from the number of respondents to the survey apart from the potential bias that might arise from differential non-response. Firstly, there are some fields of study in which the numbers are small. Secondly, even in fields of study with larger total numbers of respondents, the numbers from particular institutions may be too small to provide a reliable estimate of the learning environment. The numbers involved necessarily limit the range of analyses that can be conducted on the basis of one annual survey. It may be possible to extend the scope of analyses, if there is a cumulation of data from successive surveys.

---

## 1.4 About this report

This report of the 1999 postgraduate research experience survey is structured around six chapters. This first chapter introduces the background to the report. It outlines the nature of the data on which the analyses in the report are based. The second chapter provides a national overview of the survey results. It provides information about the distribution of item responses, information about summary statistics to characterise items and scales and

some information about the structure of the instrument. The third chapter describes some of the associations between PREQ measures and background characteristics of respondents. Chapter 4 probes differences between masters and doctoral graduates and variations among broad fields of study. Differences between masters and doctoral graduates are in the directions expected but differences among fields of study provoke speculation as to sources of difference.

Table 1: Characteristics of respondents to PREQ 1999 survey and postgraduate completions in 1998

Characteristic	Higher Education Data Collection:1998 Completions	PREQ 1999 Survey
Level of study		
Doctorate	67%	69%
Research Masters	32%	30%
Broad Field of Study		
Agriculture	4.2%	4.1%
Architecture	1.4%	1.5%
Arts	23.8%	26.2%
Business	6.6%	5.3%
Education	7.6%	9.1%
Engineering	13%	9.4%
Health	12.9%	13.3%
Law	1.2%	0.9%
Science	28.1%	29.7%
Veterinary Science	1.1%	0.5%
Gender		
Male	57.9%	53.6%
Female	42.1%	46.4%
Age		
25 and under	7.7%	2.1%
26 to 40	64.9%	62%
41 to 60	25.9%	32.8%
61 and over	1.5%	3%
Mode		
Full-time	57%	56.8%
Part-time	34.2%	34.9%
Full-time external	8.8%	7.4%

Chapter 5 contains an exploration of the patterns of difference across institutions. The extent to which there are identifiable differences among institutions is a controversial issue (Marsh, Rowe & Martin, in press). In many senses it is an issue that will become clearer when there have been successive collections of data so that the extent to which patterns are replicated can be assessed. It is an issue that is separate from the properties of the PREQ as an instrument for obtaining graduates' views of the postgraduate research experience. It would be possible for an instrument to be highly reliable in terms of generating scores reflecting the views of individual respondents but

for no differences among institutions to be detected. The final chapter draws together the results in a brief conclusion about the future uses of the PREQ and similar instruments.

## 2. National results from the PREQ

One purpose of the PREQ is to map overall patterns for postgraduates across Australia. In essence this means presenting descriptive statistics based on the opinions of the respondents to the 1999 survey. This chapter presents those descriptive statistics in several formats. It also presents information about the extent to which the views of postgraduates differ between fields of study.

### 2.1 Responses to items

#### 2.1.1 Distributions of responses

Table 2 summarises the responses of postgraduates to each item in the PREQ. It contains the wording of each of the items on the questionnaire, together with the percentages responding to each category. Respondents were asked to circle a number, between one and five, next to each item where '1' represented strong disagreement and '5' was associated with strong agreement.

As an example of the data contained in Table 2 consider the most general item: *Overall, I was satisfied with the quality of this course*. The data indicate that 3 per cent of respondents strongly disagreed with this statement and 33 per cent strongly agreed with it. The percentages responding with the intervening categories from the disagreement (2) to the agreement (4) points of the scale were 4, 10 and 51 per cent respectively. Although the intervening response points were not labelled on the questionnaire it would be reasonable to interpret them as disagree (or somewhat disagree), neutral (neither disagree nor agree) and agree (or somewhat agree). Thus, by combining the two agreement categories, it may be concluded that 84 per cent of postgraduate respondents agreed with the expression of overall satisfaction with the quality of their course. By way of comparison this indicates a slightly lower level of satisfaction than is found for the same item on the CEQ by bachelor degree graduates in 1998. From these frequency distributions it can be observed that on every item there is a greater proportion of respondents registering satisfaction (by agreeing with the item) than dissatisfaction. The balance of opinion towards agreement is stronger in some items than others. For example, there was more agreement with item 6 *my research further developed my problem-solving skills* than with item 16 *the department provided opportunities for me to become involved in the broader research culture*.

Table 2 also records the number of respondents to each item. Non-response tended to be higher on those items concerned with Infrastructure support. This may be because some postgraduate research students do not see infrastructure support as relevant, a view that may reflect the nature of their studies. On the other hand, non-response was low on items concerned with Supervision, Skill Development, Goals and Expectations, Thesis Examination and Overall Satisfaction.

**Table 2: Distribution of responses and descriptive statistics for PREQ 1999 Items: respondents from Australian universities**

Scale	Item	Strongly Disagree to Strongly Agree (%)					X <sup>a</sup>	SD	N
		SD(1)	D(2)	N(3)	A(4)	SA(5)			
<b>Supervision</b>									
1.	Supervision was available when I needed it	2.6	7.4	7.1	41.3	41.6	55.9	50.1	2196
7.	My supervisor/s made a real effort to understand difficulties I faced	4.3	7.1	14.2	36.0	38.4	48.6	54.6	2185
13.	My supervisor/s provided additional information relevant to my topic	4.7	9.5	16.0	40.5	29.3	40.1	55.2	2186
17.	I was given good guidance in topic selection and refinement	5.7	13.4	24.1	36.7	20.0	26.0	56.2	2141
21.	My supervisor/s provided helpful feedback on my progress	4.1	7.3	12.2	38.0	38.3	49.5	54.1	2194
24.	I received good guidance in my literature search	5.5	16.2	24.5	36.4	17.3	21.8	55.9	2148
<b>Skill Development</b>									
6.	My research further developed my problem-solving skills	0.7	1.2	4.9	42.6	50.6	70.6	35.2	2194
10.	I learned to develop my ideas and present them in my written work	0.6	1.7	5.8	49.5	42.5	65.8	35.5	2202
14.	My research sharpened my analytic skills	0.7	0.8	4.9	44.0	49.6	70.5	34.2	2196
20.	Doing my research helped me to develop my ability to plan my own work	0.9	1.9	9.8	46.9	40.5	62.1	38.7	2194
26.	As a result of my research, I feel confident about tackling unfamiliar problems	1.1	2.3	12.0	48.9	35.8	8.0	40.0	2181
<b>Intellectual Climate</b>									
5.	The department provided opportunities for social contact with other postgraduate students	5.0	11.6	18.5	42.6	22.3	32.9	54.8	2111
9.	I was integrated into the department's community	6.7	14.5	22.0	36.5	20.3	24.5	58.1	2057
16.	The department provided opportunities for me to become involved in the broader research culture	6.6	16.5	24.8	36.3	15.9	19.2	56.5	2102
22.	A good seminar programme for postgraduate students was provided	6.6	15.2	23.1	36.6	18.5	22.6	57.4	2115
23.	The research ambience in the department or faculty stimulated my work	8.1	6.2	29.7	31.7	14.3	13.9	56.9	2071
<b>Infrastructure</b>									
3.	I had access to a suitable working space	4.5	8.8	11.3	40.5	34.9	46.2	55.1	2014
8.	I had good access to the technical support I needed	4.0	12.5	19.0	43.8	20.8	32.5	53.2	2068
12.	I was able to organise good access to necessary equipment	2.3	7.5	17.4	48.4	24.4	42.5	47.5	1993
18.	I had good access to computing facilities and services	4.7	12.4	17.6	38.8	26.6	35.1	56.3	2015
27.	There was appropriate financial support for research activities	8.5	13.9	20.5	37.1	20.0	23.1	59.9	2065
<b>Thesis Examination</b>									
2.	The thesis examination process was fair	2.4	3.7	5.9	48.6	39.5	59.6	44.1	2185
15.	I was satisfied with the thesis examination process	5.4	6.7	7.8	44.7	35.3	48.9	54.6	2184
25.	The examination of my thesis was completed in a reasonable time	12.6	13.9	10.5	36.6	26.3	25.1	67.3	2202
<b>Goals and Expectations</b>									
4.	I developed an understanding of the standard of work expected	1.0	3.2	6.7	53.8	35.4	59.7	38.6	2191
11.	I understood the required standard for the thesis	1.3	4.1	7.7	50.9	36.0	58.0	41.7	2196
19.	I understood the requirements of thesis examination	1.1	3.8	7.8	53.6	33.7	57.5	40.1	2183
<b>Overall Satisfaction</b>									
28.	Overall, I was satisfied with the quality of my higher degree research experience	2.8	4.0	9.7	50.5	32.9	53.3	45.7	2202

<sup>a</sup> means and Standard deviations are calculated after re-coding of responses after the CEQ method, -100, -50, 0, 50, 100.

## 2.1.2 Summary statistics

Rather than present the full distribution of responses to each item it is possible to use a summary statistic. Summary statistics provide a parsimonious way of presenting data especially when comparing groups.

### Means and standard deviations

One summary statistic is the mean rating for the item<sup>1</sup>. Mean ratings for items on the PREQ are recorded in Table 2. The mean ratings in Table 2 are based on the linear transformation that has been used in the CEQ reports. In other words means are presented on a scale different from the 1, 2, 3, 4, 5 used in the questionnaire: 1 has been recoded to -100, 2 to -50, 3 to zero, 4 to +50, and 5 to +100. Consistency of interpretation of the mean scores of all items is facilitated by this approach. Negative values reflect disagreement, and positive means indicate agreement, with the meaning of the five scales formed from the individual items. As an example of the interpretation of mean scores consider item 21: *My supervisor/s provided helpful feedback on my progress*. The mean was 50. This corresponds to an average response of agree. Expressed differently it is approximately one standard deviation unit above the 'neutral, ambivalent or indifferent' category reflecting significantly more agreement than disagreement with the statement. From the distribution of the responses it can be seen that more than three-quarters of the respondents either agreed or strongly agreed with this item.

The standard deviation provides an indication of the dispersion of ratings around the mean. A larger value indicates that respondents held a wider range of views about the issue and a smaller value indicates that there was greater similarity in the views of postgraduates. In Table 2 it is interesting to note that the dispersion is rather less on some items than others. Dispersion was greatest on item 25 *the examination of my thesis was completed in a reasonable time* and least on item 14 *my research sharpened my analytic skills*. In general there was less variation in postgraduate responses to items concerned with Skill Development and with Goals and Expectations than on other items.

### Percentage agreement

A second way of summarising item response data is to reduce the five categories on the scale to just two: in agreement or not in agreement. To do

---

1 Strictly speaking the response scale is ordinal (ie one cannot assume that the differences between points on the response scale represent equal differences in level of agreement) but in using these statistics one is treating the scale as if the intervals were of equal magnitude. This approximation is commonly made in survey statistics.

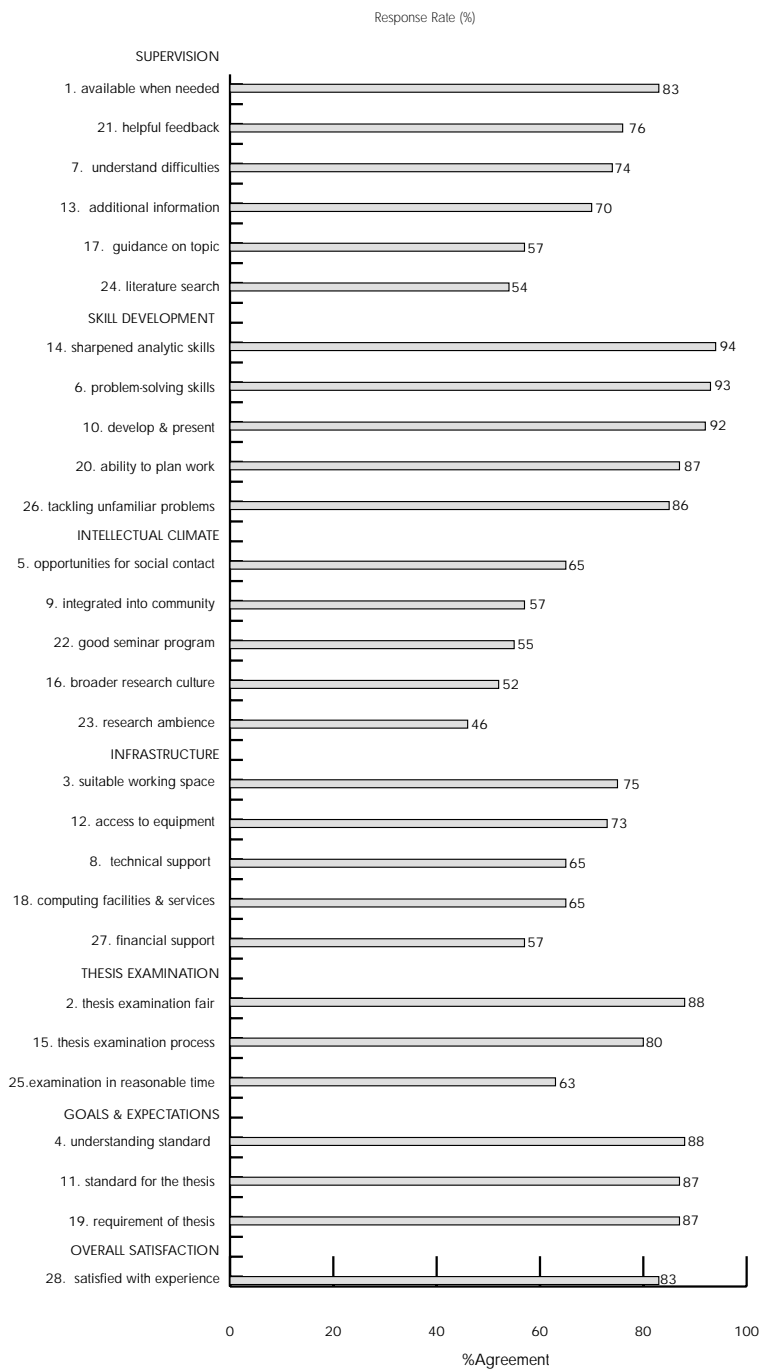
this responses 4 and 5 (agree and strongly agree) are classified as in agreement and responses 1, 2 and 3 (strongly disagree, disagree and uncertain) are classified as not in agreement. This technique reduces some of the information available in the full distribution but makes no assumption about the properties of the response scale. It allows the responses for groups of respondents to be summarised as percentage agreement that can be interpreted in relation to the item and the nature of the responses.

Table 3 records the percentage agreement with each item of the postgraduate research experience questionnaire. The same data have been represented in Figure 2. Within each group the items have been ordered by the percentage agreement.

**Table 3: Percentage agreement with items on the PREQ  
1999: Respondents from Australian Universities**

Scale	Item	Percentage Agreement
<b>Supervision</b>		
1.	Supervision was available when I needed it	83
21.	My supervisor/s provided helpful feedback on my progress	76
7.	My supervisor/s made a real effort to understand difficulties I faced	74
13.	My supervisor/s provided additional information relevant to my topic	70
17.	I was given good guidance in topic selection and refinement	57
24.	I received good guidance in my literature search	54
<b>Skill Development</b>		
14.	My research sharpened my analytic skills	94
6.	My research further developed my problem-solving skills	93
10.	I learned to develop my ideas and present them in my written work	92
20.	Doing my research helped me to develop my ability to plan my own work	87
26.	As a result of my research, I feel confident about tackling unfamiliar problems	85
<b>Intellectual Climate</b>		
5.	The department provided opportunities for social contact with other postgraduate students	65
9.	I was integrated into the department's community	57
22.	A good seminar programme for postgraduate students was provided	55
16.	The department provided opportunities for me to become involved in the broader research culture	52
23.	The research ambience in the department of faculty stimulated my work	46
<b>Infrastructure</b>		
3.	I had access to a suitable working space	75
12.	I was able to organise good access to necessary equipment	73
18.	I had good access to computing facilities and services	65
8.	I had good access to the technical support I needed	65
27.	There was appropriate financial support for research activities	57
<b>Thesis Examination</b>		
2.	The thesis examination process was fair	88
15.	I was satisfied with the thesis examination process	80
25.	The examination of my thesis was completed in a reasonable time	63
<b>Goals and Expectations</b>		
4.	I developed an understanding of the standard of work expected	89
19.	I understood the requirements of thesis examination	87
11.	I understood the required standard for the thesis	87
<b>Overall Satisfaction</b>		
28.	Overall, I was satisfied with the quality of my higher degree research experience	83

Figure 2: Percentage agreement for PREQ items



Across all items the mean percentage agreement is 73 per cent. In other words, on average, just less than three quarters of the respondents either agreed or strongly agreed with the items in the questionnaire. There were differences among the items in the extent to which respondents expressed agreement. There was a high level of agreement with the items concerned with Skill Development and with Goals and Expectations. Somewhat lower levels of agreement were expressed with items concerned with Thesis Examination, Supervision and Infrastructure. The lowest levels of agreement were recorded with the items concerned with Intellectual Climate.

Within each group of items there were differences in the percentage agreement that was reported. Among the Supervision items 83 per cent expressed agreement with the item that *supervision was available when I needed it* but only 54 per cent expressed agreement with the statement that *they received good guidance in my literature search*.

Among the items concerned with Intellectual Climate 65 per cent recorded agreement with the item *the department provided opportunities for social contact with other postgraduate students*. However, only 46 per cent recorded agreement with the statement that *the research ambience in the department or faculty stimulated my work*.

The items concerned with Thesis Examination illustrate the value of considering individual items. Even though there were high levels of agreement with the statements concerned with fairness and process there was rather less agreement with the statement that the process was completed in a reasonable time.

Overall more than four out five (83 per cent) postgraduate respondents expressed agreement with the statement *that overall, I was satisfied with the quality of my higher degree research experience*.

---

## 2.2 Groups of items and scales

The PREQ is not just a compilation of 28 items. The items form groups of related items that represent underlying dimensions of the postgraduate research experience. For example, the six items that are concerned with Supervision are of interest not only for the particular issues that they address but because they provide an indication of an underlying dimension: the perceived quality of supervision. Responses to each item from this group will result from the influence of the underlying dimension of Supervision, specific issues addressed by the item and error contributions that arise from factors such as differences in the way individuals interpret the wording of the item.

### 2.2.1 Summary statistics for groups of items or scales

There are two reasons for using a summary statistic to represent responses to groups of related items. The first reason is that such summaries provide parsimony in analysis and reporting. Rather than reporting on 28 different items it becomes possible to report patterns for six dimensions. Such a reduction of the data can be of considerable assistance in the process of making inferences about general trends and patterns. The second reason for using summary statistics for groups of items is that they use the relationships between items to confirm the meaning of the item and reduce the effect of any idiosyncrasies associated with particular items.

Two summary statistics groups of items are used in this report: the scale mean and the mean percentage agreement. It can be seen that they relate to the two summary statistics for items discussed in the previous section. The scale mean is useful in some forms of analysis where continuous measures are important (such as in various correlational and multivariate analyses) and percentage agreement is useful in representing differences between groups.

#### Scale means

Scale means are the average of the item ratings for the group of items making up the scale. The scale means provide the most reliable indicator for the group of items because they make use of the full distribution of responses to each item. It is possible to introduce greater precision by weighting each item in the scale according to the extent to which it contributes to the underlying dimension. This has not been done because the weights would be different for each set of data that was used.

#### Mean percentage agreement

Mean percentage agreement refers to the average across the items in a group of the percentage of respondents in a group agreeing or strongly agreeing with the item. It follows that this is computed for groups of respondents rather than for individual respondents. The corresponding measure for an individual would be the number of items with which they were in agreement. This measure at individual level is less reliable than the scale means, especially for scales containing only a few items, because information about the distribution of responses has not been utilised.

### 2.2.2 Scale summary statistics for the PREQ

Table 4 records two summary statistics for scales for respondents to the 1999 PREQ survey. It can be seen that the values for mean percentage agreement

match closely the values for the mean scale scores. In fact the correlation coefficient between the two sets of values in the 1999 survey data is 0.99. In a report on the CEQ, Johnson (1998) reports high correlations of approximately 0.9 between mean CEQ scale scores and the agreement percentage for the corresponding scale using data for fields of study from 1993 to 1997. From Table 4 it can also be seen that the transformation of the responses so that the range is from -100 to +100 results in mean scale scores that are of a similar order of magnitude to the figures for mean percentage agreement (especially in terms of the dispersion of the values).

Table 4: Summary statistics for PREQ sub-scales for all respondents

Scale	Number of Items	Percentage Agreement		Scale Score	
		Mean	Standard Deviation	Mean	Standard Deviation
Supervision	6	69.1	33.3	40.5	45.1
Skill Development	5	90.2	20.7	65.4	29.6
Intellectual Climate	5	54.8	35.6	22.7	45.3
Infrastructure	5	66.9	33.5	35.8	43.0
Thesis Examination	3	76.9	32.0	44.4	47.3
Goals And Expectations	3	87.7	27.0	58.4	35.6
<i>Overall Satisfaction Item</i>		83.4	37.2	53.3	45.7

It can be seen from Table 4 that respondents to the PREQ in 1999 agreed at a high rate with statements on the Skill Development and Goals and Expectations scales, such as *my research further developed my problem-solving skills* and *I developed an understanding of the standard of work required*. Items such as *the department provided opportunities for me to become involved in the broader researcher culture* on the Intellectual Climate scale were endorsed at a lower rate.

## 2.3 Structure

The responses to the 1999 PREQ were subjected to both an exploratory and a confirmatory analysis, to replicate the structure found in the developmental studies.

### 2.3.1 Exploratory principal components analysis

An exploratory principal components analysis was conducted to determine whether the structure identified in the development study was replicated in the 1999 data. Table 5 shows the pattern of factor loadings from exploratory principal components analysis using those data. The structure of, and the existence of, groupings of items in the PREQ relating to common underlying dimensions were established in the development report by ACER (1999). In accordance with convention, factor loadings less than 0.30 have been

omitted. The first column of Table 5 shows the group or scale to which the item was assigned on the basis of the developmental studies. A principal component extraction followed by an oblique (promax) rotation was used because it was hypothesised that the scales could be non-orthogonal<sup>2</sup>. Six factors had eigen values greater than one. The six factors extracted accounted for 69 per cent of the variance in item responses, and the pattern of factor loadings matched the hypothesised subscales.

Table 5: Exploratory principal components analysis of the 1999 PREQ items

Presumed Sub-scale	Item	Factor Loading					
		1	2	3	4	5	6
Supervision	1. Supervision was available when I needed it	84					
Supervision	7. My supervisor/s made a real effort to understand difficulties I faced	88					
Supervision	13. My supervisor/s provided additional information relevant to my topic	84					
Supervision	17. I was given good guidance in topic selection and refinement	75					
Supervision	21. My supervisor/s provided helpful feedback on my progress	89					
Supervision	24. I received good guidance in my literature search	74					
Skill Devpt	6. My research further developed my problem-solving skills		81				
Skill Devpt	10. I learned to develop my ideas and present them in my written work		65				
Skill Devpt	14. My research sharpened my analytic skills		87				
Skill Devpt	20. Doing my research helped me to develop my ability to plan my own work		82				
Skill Devpt	26. As a result of my research, I feel confident about tackling unfamiliar problems		83				
Infrastructure	3. I had access to a suitable working space			80			
Infrastructure	8. I had good access to the technical support I needed			73			
Infrastructure	12. I was able to organise good access to necessary equipment			85			
Infrastructure	18. I had good access to computing facilities and services			86			
Infrastructure	27. There was appropriate financial support for research activities			68			
Int. Climate	5. The department provided opportunities for social contact with other postgraduate students				84		
Int. Climate	9. I was integrated into the department's community				78		
Int. Climate	16. The department provided opportunities for me to become involved in the broader research culture				72		
Int. Climate	22. A good seminar programme for postgraduate student was provided				81		
Int. Climate	23. The research ambience in the department or faculty stimulated my work				77		
Thesis	2. The thesis examination process was fair					79	
Thesis	15. I was satisfied with the thesis examination process					90	
Thesis	25. The examination of my thesis was completed in a reasonable time					83	
Expectations	4. I developed an understanding of the standard of work expected						86
Expectations	11. I understood the required standard for the thesis						94
Expectations	19. I understood the requirements of thesis examination						81

Note: Analysis based on a principal components analysis with a promax (non-orthogonal) rotation. Decimal points have been dropped and loadings less than 0.3 have been omitted.

2 An orthogonal (varimax) rotation provided a similar result with only a slight overlap of items concerned with goals and with thesis examination

### 2.3.2 Confirmatory factor analysis

A multi-factor confirmatory factor analysis<sup>3</sup> was conducted in order to investigate the extent to which measures based on the six clusters of items identified separately identifiable constructs. Confirmatory factor analysis (CFA) differs from exploratory factor analysis (EFA) in that CFA leads to a single 'identified' solution that can be tested for 'goodness-of-fit' against the data. A CFA thus tests a 'measurement theory' against an available data set and assesses the goodness of fit of the measurement theory to the observed data.

The first analysis that was carried out specified that each item was associated with one and only one latent variable or factor. All other factor loadings were fixed to zero as were all correlations among the 'errors' of the items. All factors were allowed to be correlated. A variety of 'goodness-of-fit' measures<sup>4</sup> are recorded in Table 6. The coefficients associated with each of the items are shown in Table 7. These results indicate that the six-factor solution for the PREQ is quite strongly confirmed, and that each set of PREQ items is very largely measuring only single constructs.

Table 6: Fit indexes for multi-factor confirmatory factor analysis of the PREQ

Fit Index	Value
Chi-square	1026.9; 309 d.f.
Normed Fit Index	(NFI) 0.984
Non-normed Fit Index	(NNFI) 0.987
Comparative Fit Index (CFI)	0.989
LISREL GFI Fit Index	0.984
LISREL AGFI Fit Index	0.980
Root Mean Squared Residual (RMR)	0.101
Standardised RMR	0.101
Root Mean Square Error of Approximation (RMSEA)	0.038
90% Confidence Interval of RMSEA	(0.035, 0.040)

- 3 Dr Gerald Elsworth of the University of Melbourne conducted this part of the analysis. The confirmatory factor analyses were carried out with the structural equation modelling (SEM) programme LISREL.
- 4 In analyses with large numbers of factors and items, and a moderately large sample such as this, chi-square is regarded as an index that is excessively sensitive to lack of fit, however. A number of 'comparative fit' (or 'lack-of-fit) indices have accordingly been developed which, in various ways, compare the chi-square of the fitted model to that of a base line or 'null' model. Hence, for the present analysis, the 'comparative fit index (CFI) was 0.989 while the 'root mean square residual' (RMSR) was 0.101 and the 'root mean square error of approximation' (RMSEA) was 0.038. The first two indices can be thought of as proportional measures of goodness of fit (maximum 1.0) while the latter two can be thought of as measures of 'lack of fit'. Values of standard goodness of fit indices of over 0.9 are frequently regarded as satisfactory as are values of RMSEA of below 0.05. Hence there was a close fit of the model to the observed data.

Table 7: Coefficients from confirmatory factor analysis of the 1999 PREQ items

Assigned Sub-Scale	Item	Coefficients	1	2	3	4	5	6
Supervision	1. Supervision was available when I needed it		84					
Supervision	7. My supervisor/s made a real effort to understand difficulties I faced		91					
Supervision	13. My supervisor/s provided additional information relevant to my topic		86					
Supervision	17. I was given good guidance in topic selection and refinement		84					
Supervision	21. My supervisor/s provided helpful feedback on my progress		91					
Supervision	24. I received good guidance in my literature search		81					
Skill Devpt	6. My research further developed my problem-solving skills			85				
Skill Devpt	10. I learned to develop my ideas and present them in my written work			87				
Skill Devpt	14. My research sharpened my analytic skills			86				
Skill Devpt	20. Doing my research helped me to develop my ability to plan my own work			83				
Skill Devpt	26. As a result of my research, I feel confident about tackling unfamiliar problems			79				
Int. Climate	5. The department provided opportunities for social contact with other postgraduate students				78			
Int. Climate	9. I was integrated into the department's community				84			
Int. Climate	16. The department provided opportunities for me to become involved in the broader research culture				82			
Int. Climate	22. A good seminar programme for postgraduate student was provided				76			
Int. Climate	23. The research ambience in the department or faculty stimulated my work				83			
Infrastructure	3. I had access to a suitable working space					80		
Infrastructure	8. I had good access to the technical support I needed					85		
Infrastructure	12. I was able to organise good access to necessary equipment					85		
Infrastructure	18. I had good access to computing facilities and services					83		
Infrastructure	27. There was appropriate financial support for research activities					67		
Thesis	2. The thesis examination process was fair						89	
Thesis	15. I was satisfied with the thesis examination process						97	
Thesis	25. The examination of my thesis was completed in a reasonable time						72	
Expectations	4. I developed an understanding of the standard of work expected							91
Expectations	11. I understood the required standard for the thesis							94
Expectations	19. I understood the requirements of thesis examination							93

Note: Decimal points have been dropped.

These results indicate that the six-factor solution for the PREQ is quite strongly confirmed, and that the PREQ items are very largely measuring only single constructs. The correlations among the six PREQ latent variables suggested that there may be a second order 'positive evaluation of the research student experience' factor<sup>5</sup>. This model provided a good fit to the data. Overall, there was support for both the use of sub-scale scores and for the calculation of a total 'quality of research experience' score.

### 2.3.3 Scale reliability

Table 8 shows two measures at the individual respondent level of the reliability of the six scales formed from the 27 items of the PREQ. One of these is coefficient alpha, possibly the most commonly used measure of

5 The coefficients indicating the loading of each scale on the second order factor (quality) were as follows: Supervision (0.86), Skills (0.86), Intellectual Climate (0.86), Infrastructure (0.84), Thesis (0.72), Goals and Expectations (0.88).

reliability in social research. In fact coefficient alpha provides a lower bound estimate of the reliability of a scale. Although it has been argued that values of alpha can be inflated when there are large numbers of items in a scale, the scales in this case are rather short. Table 8 also records the values of coefficient alpha at the institutional level of analysis (ie. where institution means for each item have been used as the basis for reliability estimates).

Table 8: Reliability estimates of 1999 PREQ scales

Scale Name	No. of Items	N <sup>respondents</sup>	Composite Scale Reliability <sup>a</sup>	Cronbach Alpha Individual	Cronbach Alpha Institution
Supervision	6	2199	0.94	0.91	0.89
Intellectual Climate	5	2214	0.91	0.87	0.93
Skill Development	5	2211	0.88	0.85	0.93
Infrastructure	5	2203	0.89	0.85	0.83
Thesis	3	2217	0.88	0.80	0.71
Goals & Expectations	3	2213	0.92	0.87	0.86

<sup>a</sup> R: based on the method of Raykov (1998).

The other index of reliability in Table 8 is called the ‘composite scale reliability’ (Rc) calculated according to the method suggested by Raykov (1998). The values recorded are based on the assumption that the response options form a scale with equal intervals. It is possible to estimate the reliability when there is differential weighting of items but values from such calculations overestimate reliability when the weighting of items is not the same. Of course, measures of the internal consistency cannot be applied to a single item such as the Overall Satisfaction item.

Each of these coefficients can have values ranging between zero and unity. A higher value indicates greater internal consistency and greater reliability. Elsworth (1999) notes that an important issue in examining scales such as these is achieving balance between the breadth of coverage of the domain being measured and consistency in responses to the items that tap that domain. Items that are highly consistent with each other might only represent a very specific construct domain of little practical utility. The values of the alpha and the composite reliability coefficients in Table 8 indicate that the scales are internally consistent. All estimates of reliability of the scales are above the typically recommended ‘threshold’ values without being excessively high.

### 2.3.4 Correlations between scales

The scales (calculated from item means) are correlated with each other but not to the extent that one would presume they are simply measuring the same thing. The correlation coefficients are recorded in Table 9. The pattern is broadly similar to that reported in the development report.

Table 9: Correlation coefficients between PREQ scale scores

	Climate	Skills	Infrastructure	Thesis	Goals	Overall
Supervision	52	48	43	37	54	71
Intellectual Climate		41	58	30	41	54
Skill Development			38	34	57	54
Infrastructure				26	43	50
Thesis					45	45
Goals & Expectations						58

## 2.4 Summary

Responses from the 1999 PREQ indicate the range of responses by postgraduates to issues about their research student experience. On balance graduates expressed satisfaction with their experience but the extent of satisfaction varied among the domains of the instrument. Overall four fifths of the postgraduates were satisfied with their experience as a research student. In terms of specific areas postgraduates expressed a very high level of satisfaction with their own development of research skills but rather less with the extent to which they had been made part of the intellectual culture of their department and field. Analyses conducted of the structure of the instrument indicate that its properties in the field are congruent with those established in the development process. The principal components analysis based on the 1999 data provided results that were quite similar to those reported in the development process. Confirmatory factor analysis indicated that the model was a good fit to the data but that there may also have been a single second order factor explaining the relations between the scales. Finally the scales appeared to be internally consistent at both the individual and institutional level.

It is important to recognise the multi-factorial nature of the instrument. Although it may be possible to identify and use an overall PREQ score (especially at the aggregate level) it is more fruitful (and more conservative) to use the separate scale scores. This is equivalent to examining student achievement separately in English, mathematics, science, and studies of society and environment etc. rather than examining an overall measure of achievement.

## 3. Background characteristics of PREQ respondents in 1999

One of the important issues in the use of a survey questionnaire such as the PREQ is the extent to which responses reflect the background characteristics of respondents. It might be expected that there would be relatively few effects of background characteristics except for where differences in background were associated with the nature of the course. Ideally one would want such a questionnaire to reflect characteristics of courses of study and institutions rather than background characteristics. To the extent that this is not the case, information about the effects of background characteristics provides an important context for interpreting other effects.

This chapter will consider a selection of respondent background characteristics and their association with the scales of the PREQ. These student characteristics include gender, age at graduation, language background of the student (as well as nationality or resident status) and the mode of attendance for the course. For most of the analyses in this chapter the focus is on mean percentage agreement as the indicator of the PREQ responses. In Chapter 4 consideration is given to characteristics of the course: level of degree (masters or doctorate) and field of study.

---

### 3.1 Gender

More than 53 per cent of the graduates who responded to the PREQ in 1999 were male. This corresponds with the responses to the PREQ pilot study. In addition there were no significant gender differences in response rates for the PREQ; as had been the case for the pilot. In Table 10 mean percentage agreement for male and female respondents to the PREQ are shown, allowing a comparison of percentage agreement by gender.

There were significant differences between males and females in responses to the Infrastructure items (where the difference is more than 9 percentage points) and Overall Satisfaction (where difference is just a little less than 4 percentage points)<sup>6</sup>. In both cases the direction of the difference was that males expressed greater satisfaction on these scales. Of course it might be that

---

<sup>6</sup> The procedure for testing the significance of the difference between these groups, and other groups reported in this chapter, took account of the clustered nature of the data.

the differences observed could be associated with differences in the fields of study in which males and females study; for example, males tending to undertake postgraduate research in the sciences where Infrastructure scores might be expected to be higher and females studying in the arts and humanities where Infrastructure scores tend to be lower.

There was a tendency for males to score higher than females on the Intellectual Climate scale (significant at the 10 per cent level) by 2.5 percentage points. Again this could be associated with concomitant differences in field of study. Differences on the Supervision, Skill Development, Thesis Examination and Goals and Expectations scales were less than 2 percentage points and not statistically significant.

Table 10: Mean percentage agreement with PREQ sub-scales for male and female respondents

Respondent Gender	Sub-Scale	Mean Percentage Agreement	Standard Deviation	N respondents to PREQ
Males	Supervision	69.8	33.4	1187
	Skill Development	89.7	21.8	1191
	Intellectual Climate	56.0	34.8	1147
	Infrastructure	71.3	31.9	1111
	Thesis Examination	77.3	31.5	1192
	Goals And Expectations	88.5	26.4	1194
	Overall Satisfaction Item	85.2	35.6	1193
Females	Supervision	68.3	33.1	1030
	Skill Development	90.9	19.3	1034
	Intellectual Climate	53.5	36.6	994
	Infrastructure	61.9	34.6	968
	Thesis Examination	76.5	32.7	1028
	Goals And Expectations	86.8	27.5	1031
	Overall Satisfaction Item	81.5	38.8	1030

### 3.2 Age of respondents

The average age for respondents to the 1999 PREQ was 38 years and six months. One quarter were 30 years of age or younger and one quarter were older than 44 years. The median age was 36 years. In line with the developmental study of the PREQ, respondents were divided into four age groups for analysis of differences based on age. These were:

- 25 years and under;
- 26 to 40 years old;
- 41 to 60 years old; and
- Over 60 years old.

Approximately 2 per cent of respondents were 25 years of age or younger, 62 per cent were in the 26 to 40 age group, approximately 33 per cent in the 41 to 60 age group, and 3 per cent were aged over 60 years at the time of the survey. Table 11 contains data on the levels of agreement with the PREQ scales by these different age groups.

Table 11: Mean percentage agreement with PREQ sub-scales by respondent age groups

Age Group	Sub-Scale	Mean Percentage Agreement	Standard Deviation	N <sup>respondents to PREQ</sup>
25 and Under	Supervision	63.7	35.6	45
	Skill Development	84.4	26.9	45
	Intellectual Climate	52.8	31.6	44
	Infrastructure	68.4	31.0	44
	Thesis Examination	85.2	26.2	45
	Goals And Expectations	88.1	27.7	45
	Overall Satisfaction Item	80.0	40.5	45
26 to 40	Supervision	68.0	33.1	1357
	Skill Development	91.9	18.4	1363
	Intellectual Climate	55.8	34.9	1331
	Infrastructure	69.5	32.1	1334
	Thesis Examination	76.4	31.9	1359
	Goals And Expectations	87.8	26.7	1362
	Overall Satisfaction Item	83.5	37.1	1363
41 to 60	Supervision	70.8	33.4	710
	Skill Development	87.4	24.4	714
	Intellectual Climate	52.9	37.0	666
	Infrastructure	61.9	35.6	618
	Thesis Examination	77.0	32.8	710
	Goals And Expectations	88.0	27.1	711
	Overall Satisfaction Item	83.5	37.2	708
Over 60	Supervision	79.0	28.6	65
	Skill Development	90.1	16.9	63
	Intellectual Climate	54.8	38.2	60
	Infrastructure	57.8	36.8	45
	Thesis Examination	81.3	30.5	66
	Goals And Expectations	87.1	24.4	67
	Overall Satisfaction Item	88.1	32.7	67

When a comparison was made between graduates age 40 years or less and those aged more than 40 years statistically significant differences were observed on the Infrastructure scale and the Skill Development scale. In terms of Infrastructure, the data indicated that older graduates were less satisfied with the provision of technical support and equipment than were younger graduates. The magnitude of the difference was approximately 8 percentage points. In terms of Skill Development the difference was smaller (around 4 percentage points) and less clear. Overall, older graduates gave lower

ratings to the items on this scale but it should be noted the effect largely resulted from the responses of the youngest graduates, those aged 25 years or less.

On the Supervision scale there was a tendency for the very youngest and oldest graduates to differ in the direction of levels of satisfaction. It is interesting to speculate about some possible reasons for this difference between the youngest and oldest graduates. It is possible that it reflects differences in the way older students approach and interact with supervisors. Alternatively it could reflect differences in the ways in which supervisors respond to older research students. Yet again, it may be that the different experience that older students bring to the task influences what they expect of, and how they make use of, supervisors. Or, it could be a reflection of how much, and what type of, supervision and assistance are expected by students of different ages. However, it should be noted that the difference between the two middle age groups was not significant.

On the other scales the differences between graduates 40 years of age or younger and those older than 40 were not statistically significant. In terms of Overall Satisfaction the responses of the oldest graduates tended to be more positive than other graduates.

---

### **3.3 Language background and resident status**

The majority of respondents to the PREQ in 1999 were Australian residents (over 87 per cent), although more than 26 per cent of the respondents indicated that they were from a non-English speaking background. This is in line with the findings of the ACER developmental report (ACER, 1999) which reported that a quarter of the developmental sample was of a non-English speaking background.

There were statistically significant differences between respondents of a non-English speaking background and their English-speaking background counterparts on the Supervision (5 points), Infrastructure (8 points), and Goals and Expectations scales (4 points) and on the Overall Satisfaction item (6 points). In addition there was a tendency for the two groups to differ on Intellectual Climate scale (by 4 points but at the 10 per cent level of significance). On all these scales the difference was in the direction of more favourable responses by postgraduates of non-English speaking background (see Table 12).

Table 12: Mean percentage agreement with PREQ sub-scales for non-English speaking background and English speaking background respondents

NESB Status	Sub-Scale	Mean Percentage Agreement	Standard Deviation	N <sup>respondents</sup> to PREQ
Non-English Speaking Background				
	Supervision	72.7	32.1	588
	Skill Development	91.5	19.7	589
	Intellectual Climate	58.1	35.2	577
	Infrastructure	72.9	30.5	574
	Thesis Examination	77.8	31.0	591
	Goals And Expectations	90.3	23.2	591
	Overall Satisfaction Item	87.8	32.8	590
English Speaking Background				
	Supervision	67.9	33.6	1626
	Skill Development	89.8	21.1	1633
	Intellectual Climate	53.7	35.8	1561
	Infrastructure	64.7	34.3	1502
	Thesis Examination	76.6	32.4	1626
	Goals And Expectations	86.8	28.1	1631
	Overall Satisfaction Item	81.9	38.5	1630

There are several interpretations that can be placed on the differences associated with language background. One could be that it reflects differences in motivation that result in different interpretations being placed on the same situation. Another is that it reflects the success of various support programmes initiated by universities to bridge language gaps and provide support for students of non-English speaking backgrounds.

Table 13 records the level of agreement for the PREQ scales for Australian resident and overseas respondents. Those data indicate more favourable ratings by overseas postgraduates on the Supervision and Intellectual Climate scales, on the Infrastructure scale (by more than 10 points) and on the Overall Satisfaction item (6 points). In general this finding suggests that Australian tertiary institutions are providing a supportive environment for students who come to Australia to pursue postgraduate research studies.

Table 13: Mean percentage agreement with PREQ sub-scales for Australian and overseas resident respondents

Resident Status	Sub-Scale	Mean Percentage Agreement	Standard Deviation	N <sup>respondents</sup> to PREQ
Australian Residents				
	Supervision	68.3	33.5	1936
	Skill Development	89.9	21.0	1944
	Intellectual Climate	53.7	35.6	1865
	Infrastructure	65.3	34.0	1806
	Thesis Examination	76.8	32.2	1939
	Goals And Expectations	87.4	27.3	1944
	Overall Satisfaction Item	82.7	37.8	1942
Overseas Residents				
	Supervision	75.3	31.4	277
	Skill Development	92.7	18.7	277
	Intellectual Climate	62.8	34.7	272
	Infrastructure	77.7	28.3	269
	Thesis Examination	77.3	31.3	277
	Goals And Expectations	89.6	24.3	277
	Overall Satisfaction Item	88.8	31.6	277

### 3.4 Mode of study

The majority of respondents to the 1999 PREQ had studied for their higher degrees full-time (65 per cent), in line with the finding in the developmental study (ACER, 1999) that 67 per cent of their sample had been enrolled to study full-time. As shown by the data recorded in Table 14, respondents who had completed their higher degree part-time agreed with items of the Intellectual Climate (by 7 percentage points) Infrastructure scales (by 12 percentage points) to a lesser extent than their full-time counterparts. These differences are statistically significant. It is understandable that students studying part-time do not feel that they have the same access to technical support and equipment as students who are studying full-time, and do not feel as much that they are part of the intellectual culture of their institution.

Table 14: Mean percentage agreement to PREQ sub-scales for full-time and part-time respondents

Mode of Study	Sub-Scale	Mean Percentage Agreement	Standard Deviation	N <sup>respondents</sup> to PREQ
Full-Time Students				
	Supervision	67.9	33.9	1449
	Skill Development	91.0	20.3	1448
	Intellectual Climate	57.3	34.8	1431
	Infrastructure	70.9	31.7	1418
	Thesis Examination	76.7	31.8	1447
	Goals And Expectations	87.8	26.6	1448
	Overall Satisfaction Item	83.1	37.5	1445
Part-Time Students				
	Supervision	71.3	32.0	765
	Skill Development	88.8	21.6	774
	Intellectual Climate	50.2	36.8	708
	Infrastructure	59.1	35.9	658
	Thesis Examination	77.2	32.6	769
	Goals And Expectations	87.5	27.6	774
	Overall Satisfaction Item	84.0	36.7	776

### 3.5 Summary

There are some differences in PREQ responses associated with differences in the background characteristics of respondents. In terms of Overall Satisfaction, males recorded more favourable views than females (a difference of 4 points), graduates older than 40 years of age recorded more favourable views than younger graduates (a difference of 4 points) and graduates of non-English recorded more favourable views than graduates of an English speaking background (by 6 points). In addition graduates who were overseas residents recorded higher levels of Overall Satisfaction than Australian graduates did (by 6 percentage points).

The biggest differences concerned Infrastructure. There was a 9-point difference between males and females (in favour of males), a 12-point difference between full-time and part-time students (in favour of full-time students), and a 9-point difference between graduates under the age of 40 and those over the age of 40 (in favour of the younger graduates). In addition there was an 8-point difference between graduates of non-English speaking background and those of English speaking background, and a 12-point difference between graduates of overseas residential status and those of Australian residential status.

There were also differences associated with background characteristics on the Supervision scale. These reflected age (a 9-point difference in favour of graduates older than 40), residential status (a 7-point difference in favour of overseas students) and non-English speaking background (a 5-point difference

in favour of non-English speaking background). In addition there was a 3 point difference in favour of graduates who were older than 40 years. On the Intellectual Climate scale those who had been full-time students recorded more favourable views than those who had been part-time students (by 7 points), overseas graduates recorded more favourable views than Australian residents (by 9 points) and graduates of non-English speaking background recorded more favourable views than other graduates (by 4 points).

The sources of these differences remain unexplored. They could arise from differences in the types of programme in which students are enrolled, differences in perceptions of the same situation or differences that arise from variations in programmes or treatment within programmes.

## 4. Differences between levels and field of study

An instrument such as the PREQ can serve a number of purposes. Arguably one of those purposes is to provide a tool that can illuminate the influences of programme and organisational factors on the quality of the postgraduate learning experience. Mapping differences between the course experiences of doctoral and masters graduates and among fields of study can provide a broad indication of some of those influences. The previous chapter was concerned with the associations between the PREQ scales and the background characteristics of respondents. It is necessary to appreciate that background factors can influence perceptions of a learning environment. The present chapter is more concerned with aspects that are associated with postgraduate programmes of study: the level of study and the field of study.

---

### 4.1 Level of study

Nearly 70 per cent of respondents to the 1999 PREQ survey were doctoral (PhD) graduates, and 30 per cent had completed a research masters degree. Response rates for respondents to the PREQ who were completing research masters degrees and PhDs are discussed in Chapter 1. Table 15 reports the percentage agreement for the PhD and research masters graduates separately.

Doctor of philosophy graduates recorded significantly higher levels of agreement with items on several scales of the PREQ than did research masters degree graduates. There were statistically significant differences between these two levels of degree on the Skill Development scale, the Intellectual Climate scale, the Infrastructure scale, the Goals and Expectations scale, and (marginally) the Overall Satisfaction item. There was no significant difference between the degree levels on the Supervision or Thesis Examination scales.

The largest difference was on the Infrastructure scale. These items dealt largely with the provision of equipment and support, both technical and financial. Respondents to the 1999 PREQ who had graduated at doctoral level agreed to a considerably greater extent (70.5 per cent) with the items on this scale than those who had graduated with a research masters (58.5 per cent): a difference of 12 percentage points.

Doctoral respondents also reported a higher rate of agreement with items on the Intellectual Climate scale. Doctor of philosophy graduates indicated that

their departments had provided them with the support and opportunity to integrate with and become involved with both the departmental and broader research communities to a greater extent than did respondents who had completed a research masters degree. The difference was 10 percentage points (58 per cent agreement as compared to 48 per cent agreement).

Table 15: Mean percentage agreement with PREQ sub-scales for research masters and doctor of philosophy respondents

Level of Study	Sub-Scale	Mean Percentage Agreement	Standard Deviation	N respondents to PREQ
Research Masters Graduates				
	Supervision	66.9	34.8	658
	Skill Development	86.2	24.0	659
	Intellectual Climate	47.8	35.9	629
	Infrastructure	58.5	34.2	609
	Thesis Examination	75.7	32.2	651
	Goals And Expectations	82.6	31.1	657
	Overall Satisfaction Item	79.2	40.6	580
PhD Graduates				
	Supervision	70.0	32.7	1549
	Skill Development	91.9	19.0	1556
	Intellectual Climate	58.0	35.1	1502
	Infrastructure	70.5	32.6	1461
	Thesis Examination	77.4	32.0	1559
	Goals And Expectations	89.8	24.8	1559
	Overall Satisfaction Item	85.2	35.5	1555

Results such as these are not surprising. One might expect that doctoral graduates would have experienced higher levels of Infrastructure Support than masters graduates. And, one would expect that doctoral graduates would have been included in the broader research community to a greater extent than masters graduates. These results also provide some evidence of the validity of the measures.

Doctor of philosophy graduates also reported that they developed a greater understanding of expectations for them than did masters graduates (the difference was 7 points) and had developed skills to a greater extent (the difference was 6 points). The Overall Satisfaction of doctoral graduates was higher than that of masters graduates (the difference was 6 points). In general the experience of a doctoral programme appears to be a richer experience than a research masters programme on many of the domains of the PREQ.

## 4.2 Broad field of study

One might expect the formal requirements and broad focus of a higher research degree in Economics, for example, would be largely similar across institutions. The development report (ACER, 1999) reported support for aggregation to the level of fields of study. Caution needs to be taken regarding the position for veterinary science and law where there are a limited number of respondents. Although data for law and veterinary science are recorded they have not been referred to in the examination of differences between fields of study. PREQ scale scores, expressed as mean percentage agreement for different fields of study, are reported in Table 16.

Table 16: Mean percentage agreement with PREQ sub-scales by broad field of study

Broad Field of Study	Sub-Scale	Mean Percentage Agreement	Standard Deviation	N <sub>respondents</sub> to PREQ
Agriculture	Supervision	67.4	35.8	91
	Skill Development	92.7	17.5	91
	Intellectual Climate	59.0	35.3	88
	Infrastructure	78.8	25.8	89
	Thesis Examination	72.9	34.4	91
	Goals And Expectations	93.0	19.6	91
	Overall Satisfaction Item	84.4	36.5	90
Architecture, Building	Supervision	58.2	39.5	33
	Skill Development	87.0	21.2	33
	Intellectual Climate	56.2	37.4	33
	Infrastructure	66.2	35.9	29
	Thesis Examination	65.7	38.6	33
	Goals And Expectations	82.8	27.8	33
	Overall Satisfaction Item	72.7	45.3	33
Arts, Humanities & Social Sciences	Supervision	69.0	34.1	579
	Skill Development	87.8	22.9	581
	Intellectual Climate	49.6	36.2	565
	Infrastructure	53.6	37.1	525
	Thesis Examination	74.4	34.2	579
	Goals And Expectations	84.2	30.3	582
	Overall Satisfaction Item	79.2	40.6	581
Business, Administration, Economics	Supervision	67.3	33.6	119
	Skill Development	90.3	23.2	119
	Intellectual Climate	52.1	39.5	113
	Infrastructure	66.4	35.9	107
	Thesis Examination	78.8	31.3	118
	Goals And Expectations	86.6	28.6	119
	Overall Satisfaction Item	84.0	36.8	119

(continued)

**Table 16:** Mean percentage agreement with PREQ sub-scales by broad field of study (*continued*)

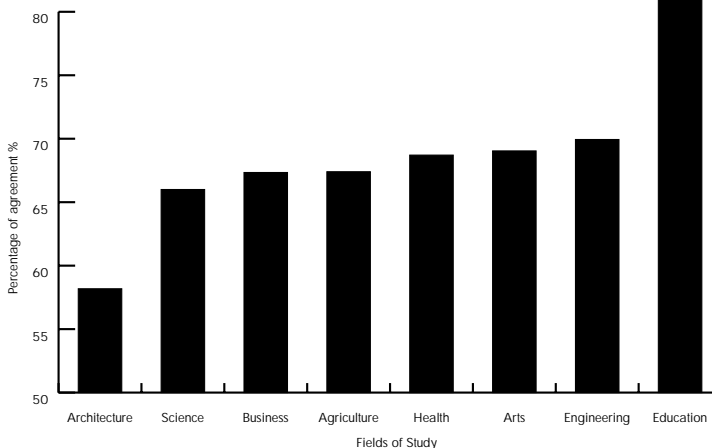
Broad Field of Study	Sub-Scale	Mean Percentage Agreement	Standard Deviation	N <sup>respondents</sup> to PREQ
Education				
	Supervision	81.7	27.4	204
	Skill Development	91.2	19.6	204
	Intellectual Climate	57.0	37.8	185
	Infrastructure	58.6	35.1	169
	Thesis Examination	82.9	28.1	202
	Goals And Expectations	91.7	22.4	202
	Overall Satisfaction Item	88.2	32.3	204
Engineering, Surveying				
	Supervision	69.9	32.4	208
	Skill Development	92.4	18.5	209
	Intellectual Climate	58.8	34.1	202
	Infrastructure	73.3	30.7	204
	Thesis Examination	74.2	33.2	209
	Goals And Expectations	89.7	24.9	208
	Overall Satisfaction Item	85.7	35.2	208
Health, Medicine, Dentistry				
	Supervision	68.7	32.0	291
	Skill Development	92.1	17.1	292
	Intellectual Climate	55.0	34.5	286
	Infrastructure	71.1	30.6	283
	Thesis Examination	78.5	31.3	293
	Goals And Expectations	87.9	27.7	293
	Overall Satisfaction Item	85.3	35.5	293
Law, Legal Studies				
	Supervision	72.8	29.0	19
	Skill Development	84.7	27.2	19
	Intellectual Climate	39.3	32.4	14
	Infrastructure	57.6	34.1	15
	Thesis Examination	78.9	29.8	19
	Goals And Expectations	86.0	25.6	19
	Overall Satisfaction Item	89.5	31.5	19
Science				
	Supervision	66.0	33.6	662
	Skill Development	90.5	20.8	666
	Intellectual Climate	57.7	34.3	643
	Infrastructure	74.6	28.7	646
	Thesis Examination	78.2	30.2	664
	Goals And Expectations	88.5	25.9	667
	Overall Satisfaction Item	84.4	36.4	665
Veterinary Science				
	Supervision	57.6	39.0	11
	Skill Development	89.1	25.9	11
	Intellectual Climate	43.6	36.7	11
	Infrastructure	67.3	32.6	11
	Thesis Examination	69.7	40.7	11
	Goals And Expectations	90.9	21.6	11
	Overall Satisfaction Item	63.6	50.5	11

There are differences between fields of study on several domains of the PREQ. Some of those differences, such as the 24-percentage point difference in Supervision ratings between education and architecture, indicate areas where further inquiry might be warranted. In the sections that follow, patterns of difference across broad fields of study are described for five of the scales and for the Overall Satisfaction item<sup>7</sup>. For Goals and Expectations, the gap between the highest and lowest mean score is a little over 11 percentage points. On that scale the differences between fields of study are not statistically significant and are not discussed in the sections that follow.

### 4.2.1 Supervision

As shown in Figure 3 on the Supervision scale, research degree graduates from Education reported higher rates of satisfaction than did graduates from most other fields of study. The difference between education and engineering (12 percentage points) is statistically significant. At the other end of the distribution the differences between architecture and each of health, the arts, engineering (around 11 percentage points) are statistically significant. In fact the gap in Supervision rating between education and architecture is 24 percentage points.

Figure 3: Mean percentage agreement for supervision by field of study

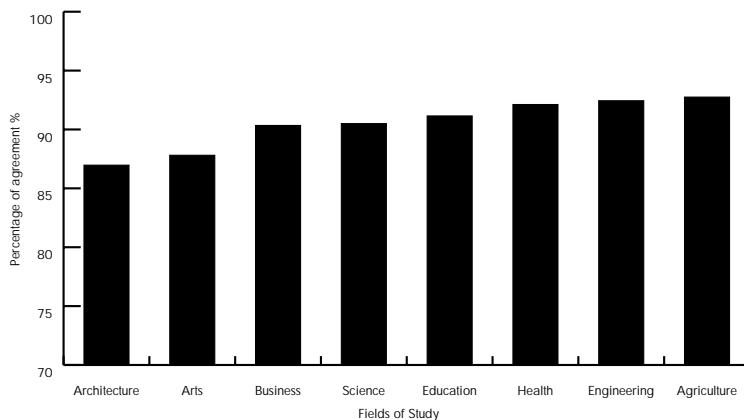


<sup>7</sup> The highest and lowest scores are associated with two small fields of study not shown: law and veterinary science.

### 4.2.2 Skill development

On the Skill Development items there was a small range in mean percentage agreement across fields of study. The range shown in Figure 4 from the lowest (both architecture and the arts) to the highest (agriculture) was 5 to 6 percentage points. Nevertheless, the fields of agriculture, engineering, health and science were significantly different from the arts, humanities and social sciences.

Figure 4: Mean percentage agreement for skill development by field of study

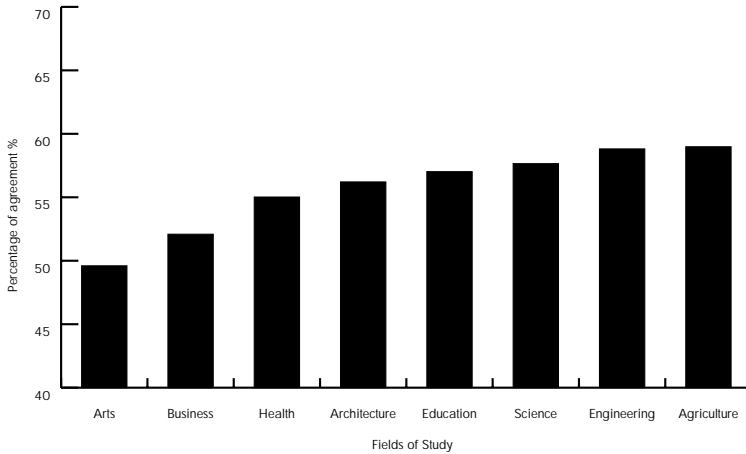


### 4.2.3 Intellectual climate

The distribution of mean percentage agreement scores by field of study is shown in Figure 5. The range from the lowest (arts) to the highest (agriculture) was 9 percentage points<sup>8</sup>. Taking the arts, humanities and social science as the reference, there were statistically significant differences between graduates in this field and agriculture, engineering, science and education. This indicates that graduates from postgraduate courses of study in the arts were less satisfied with the seminar programmes and the research culture of their departments than were graduates in these four fields of study. It is possible that in science-based fields of study it is more common for a postgraduate research student to become part of a research team or laboratory than in the arts. This would contribute to becoming more a part of the intellectual culture of the field.

<sup>8</sup> Graduates in veterinary science and in law reported the lowest agreement with items on the Intellectual Climate Scale. They have not been shown in Figure 5 because of the small numbers involved.

Figure 5: Mean percentage agreement scores for intellectual climate by field of study



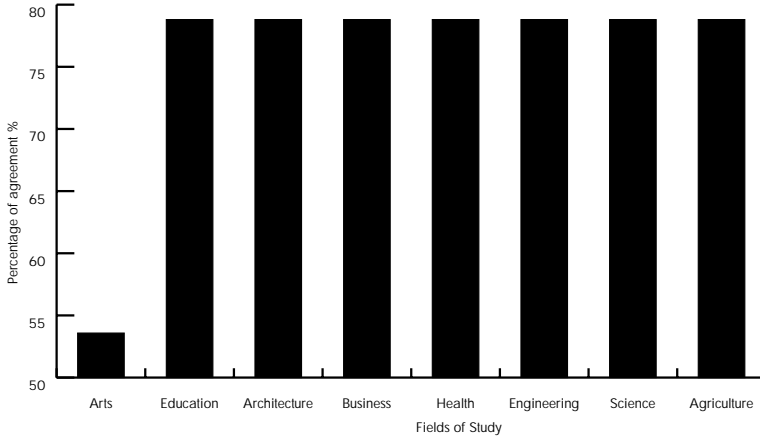
#### 4.2.4 Infrastructure

Figure 6 displays the mean percentage agreement scores on the Infrastructure scale across fields of study. These scores range from 54 per cent for the arts, humanities and social sciences to 79 per cent for agriculture, a gap of 25 percentage points. The gap between the two largest fields of study (arts and science) was 21 percentage points. All of the other fields of study differ significantly from the arts.

If education is taken as the reference field there are significant differences between it and the fields of agriculture, engineering, science and health. In addition the arts, humanities and social sciences are significantly lower than education.

It may be that technical equipment and support are more readily available in fields of study that emphasise the use of technology, such as engineering, than in areas such as languages and or history. This availability is then reflected in the response of the graduates of postgraduate research degrees. An additional possibility is that in fields where postgraduate students become part of laboratory groups that are concerned with experimental research they have greater access to a range of appropriate technical equipment and support.

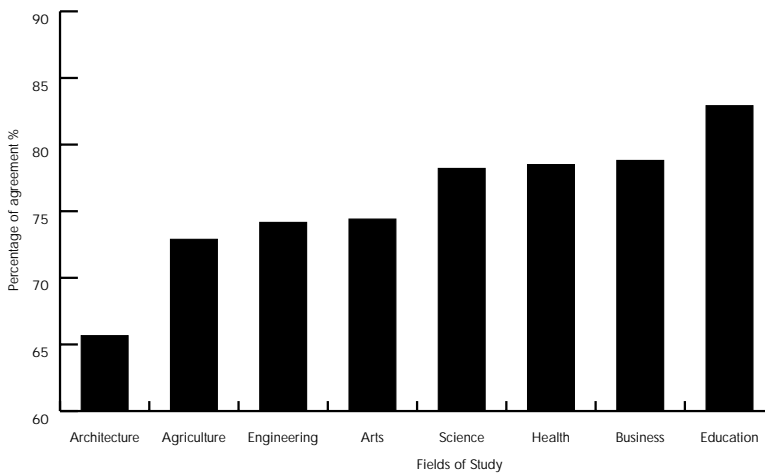
Figure 6: Mean percentage agreement for infrastructure by field of study



#### 4.2.5 Thesis examination

As shown in Figure 7 there was a range of 16 percentage points in the mean percentage agreement scores between the lowest (architecture) and highest (education) fields of study. Scores from education and business differed significantly from scores in architecture. Although there is no immediately apparent interpretation that can be offered for this pattern it is plausible that there may be different expectations regarding thesis examination procedures in different fields of study.

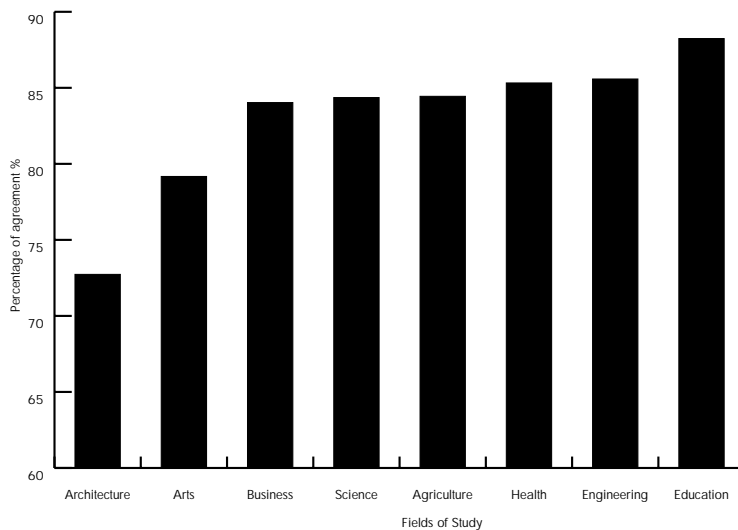
Figure 7: Mean percentage agreement for thesis examination by field of study



### 4.2.6 Overall satisfaction

Measures of Overall Satisfaction are based on responses to a single item. As a consequence there is more uncertainty in these measures than for the scales that contain five or six items. The percentage agreement scores on Overall Satisfaction range from 73 per cent (architecture) to 88 per cent (education). Results for each field of study are shown in Figure 8. Figure 8 shows a narrow spread between education, engineering, health, agriculture, science, and business with a falling away to arts and architecture. The differences between the arts (taken as the reference field) and health, engineering or science are statistically significant (and around 5 to 6 percentage points). The wider gap between education and the arts is not significant because of the greater dispersion among education scores.

Figure 8: Mean percentage agreement for overall satisfaction by field of study



## 4.3 Programme level and field of study

The preceding sections of the present chapter have investigated the association of PREQ scores with level of programme and field of study. This section concerns analyses of the interaction between level and field of study on PREQ scales.

Overall 70 per cent of respondents to the PREQ had completed doctoral programmes but this varied a little according to field of study. Table 17 records the distribution of respondents between these two types of postgraduate research programmes for each field of study.

Table 17: Percentage respondents in different programmes by field of study

Field of Study	% Respondents		Number of Respondents
	Masters	Doctorate	
Agriculture	34.4	65.6	90
Architecture	44.1	55.9	34
Humanities, Arts and Social Sciences	36.0	64.0	577
Business Studies	32.5	67.5	114
Education	32.8	67.2	198
Engineering	35.1	64.9	208
Health	30.7	69.3	293
Law	30.0	70.0	20
Science	21.1	78.9	658
Veterinary Science	20.0	80.0	10
<i>Total</i>	<i>30.2</i>	<i>69.8</i>	<i>2202</i>

Note: 23 respondents did not indicate a field of study

The lowest percentage of doctoral respondents is from architecture (56 per cent) and the highest is from Veterinary Science (80 per cent). As these two fields have relatively low total numbers of respondents it may be better to represent the range for the bulk of respondents as from Humanities, Arts and Social Sciences (64 per cent were doctoral graduates) and Science (79 per cent were doctoral graduates). Although this range is not large, the observation that there are overall differences by level of programme suggests that it may be important to consider differences among fields of study for each programme level.

Table 18 contains the scale scores (expressed as mean percentage agreement) for eight fields of study separately for masters and doctoral graduates. These data have been represented as profiles in Figure 9.

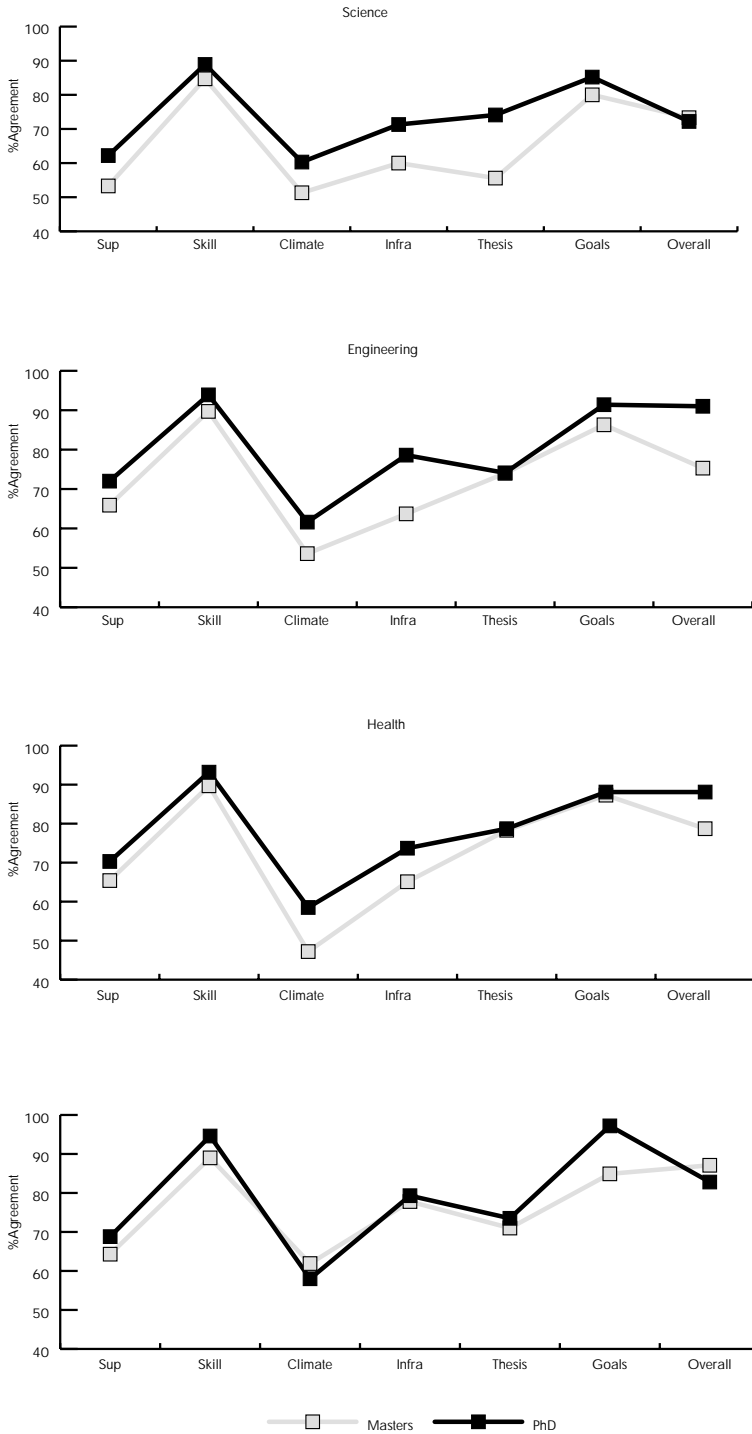
Table 18: Percentage agreement scores by level of programme and field of study

Field of Study	Level	Percentage Agreement on Scale						
		Supervision	Skill Devpt	Intell.Climate	Infrastructure	Thesis	Goals	Overall Satisfaction
Agriculture	PhD	68.8	94.6	58.0	79.3	73.5	97.2	82.8
	Masters	64.3	89.0	61.9	77.8	71.0	84.9	87.1
Architecture	PhD	62.2	88.9	60.3	71.3	74.1	85.2	72.2
	Masters	53.3	84.7	51.3	60.0	55.6	80.0	73.3
Humanities	PhD	71.4	90.3	54.2	58.3	76.6	89.2	82.0
	Masters	65.7	82.8	41.7	44.7	71.0	75.8	74.8
Business Studies	PhD	71.3	94.5	61.2	71.8	78.8	90.0	88.3
	Masters	59.3	80.3	38.4	57.4	79.6	80.2	75.7
Education	PhD	80.9	91.9	59.4	63.9	81.7	92.4	88.6
	Masters	84.4	89.2	52.7	48.8	86.5	90.4	87.7
Engineering	PhD	72.0	93.9	61.6	78.6	74.1	91.4	91.0
	Masters	65.9	89.7	53.6	63.7	74.0	86.3	75.3
Health	PhD	70.3	93.2	58.5	73.7	78.7	88.1	88.1
	Masters	65.4	89.7	47.2	65.1	78.3	87.3	78.7
Science	PhD	65.9	90.9	59.9	75.9	78.1	89.5	84.6
	Masters	65.8	88.0	51.0	69.3	78.3	84.1	83.3

Note: Entries for Law and Veterinary Science have been omitted because of the small numbers in those fields.

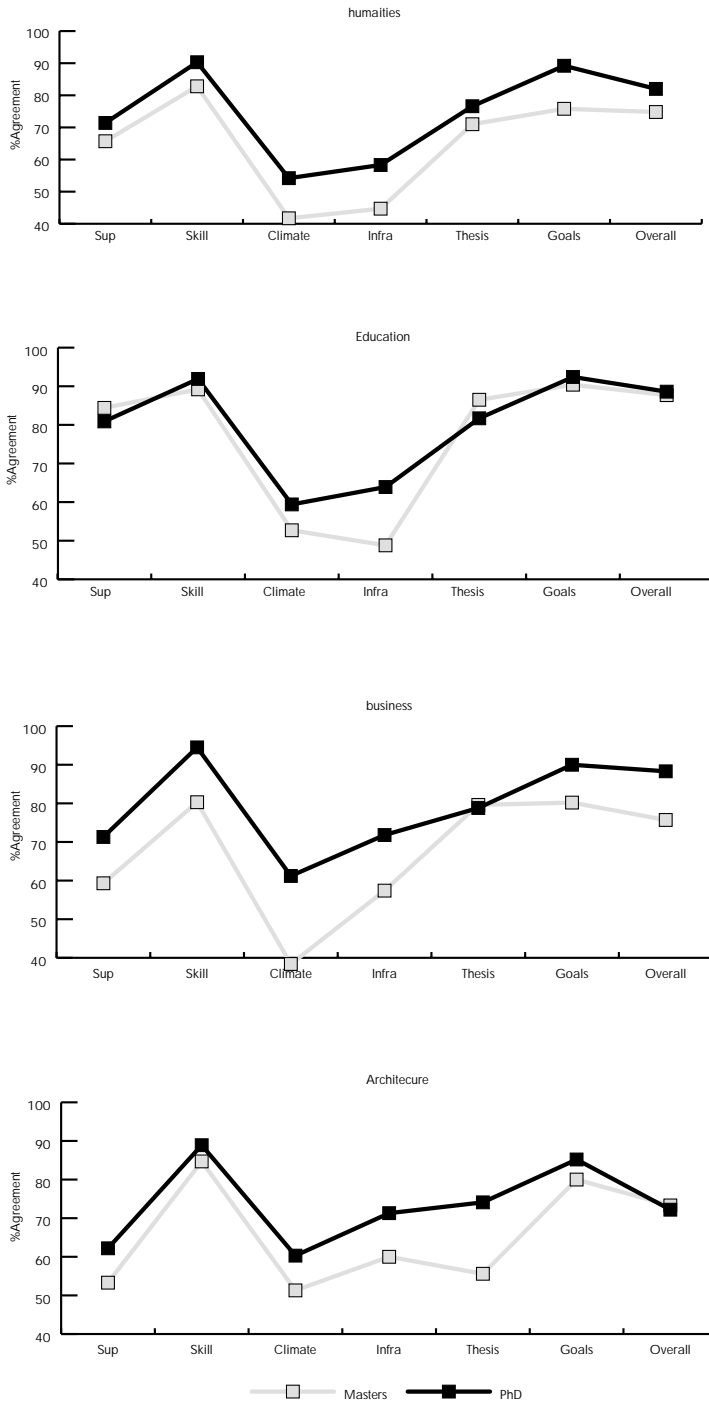
In general it can be seen that there is a broad correspondence between the profiles of masters and doctoral graduates. On average the correlation coefficient between the two sets of data for each field of study is 0.96. The match is a little lower for engineering, architecture and agriculture (where the coefficients are around 0.90) and somewhat higher for humanities, health and science (where the coefficients are around 0.97). The discrepancy between the ratings of masters and doctoral graduates is greatest for Infrastructure and Intellectual Climate and is substantial on those scales for all fields of study except agriculture. A general conclusion could be that it would be possible to combine the data for masters and doctoral graduates for some analyses. However, in the initial examination of institutional differences the two levels were kept separate.

Figure 9: Profiles of PREQ scales by programme level and field of study



(continued)

Figure 9: Profiles of PREQ scales by programme level and field of study (continued)



---

## 4.4 Summary

There were differences in postgraduates' views of their research student experience between doctoral and masters degree graduates and between fields of study. The differences between doctoral and masters' graduates were in the expected directions for each of the domains for which substantial differences were observed. On the domains of the PREQ where differences were observed it appeared that the experience of a doctoral programme was richer than the experience of a research masters programme.

Differences among fields of study were more difficult to interpret. There were moderately sized differences in Overall Satisfaction, Infrastructure, Supervision and Intellectual Climate. Those differences were not the same across the domains; a result that invites speculation about the cause and nature of those differences. One can imagine plausible explanations for the relatively low rating accorded to the Infrastructure domain by postgraduates in the arts, humanities and social sciences in terms of what is provided. It is less clear why the ratings from that field should be relatively low in terms of Intellectual Climate and Overall Satisfaction. Education postgraduates also rate their experience poorly in terms of Infrastructure support but highly in terms of Supervision and Overall Satisfaction and reasonably in terms of Intellectual Climate.

## 5. Institutional patterns

The extent to which analyses of the PREQ can, or should be, used to identify institutional differences is an important and contentious issue. It is important because providing information on postgraduates' satisfaction with their experiences could provide a basis for individual institutions to respond to areas of concern by gathering additional information at local level and introducing changes to such things as supervision, intellectual climate and infrastructure. At a broader level analysis of PREQ data by institution could also be used to identify the location of good practice (or more precisely institution and fields of study with above average scores on relevant scales of the PREQ). Both of these could be thought of as formative purposes directed towards the improvement of research training for postgraduate students. Other purposes for analysing PREQ data by institution might be to inform students who are choosing where to pursue postgraduate study and to generate indicators that could be incorporated in an accountability framework.

Linked to these potential purposes are technical issues concerning the extent to which scores reflect the policies and practices of institutions and the extent to which one can be confident that observed differences between institutions have not arisen by chance.

The first of these issues concerns the effect of influences other than policy and practice within institutions. It is widely recognised that simple summary scores for institutions could reflect differences in the mix of fields of study represented among its respondents (ACER, 1999). For example, other things being equal an institution with a preponderance of enrolments in science-based fields of study might be expected to score more highly on the Infrastructure scale. Similarly, an institution with a greater proportion of doctoral students might be expected to score higher on the Intellectual Climate scale than an institution with a greater proportion of masters students. Differences in institutional indicators could also be influenced by differences in the background characteristics (such as sex and age) of graduates, although analyses reported in Chapter 3 suggest these effects are small. Indicators that seek to reflect differences in institutional policy and practice need to have made allowance for as many of these influences as possible.

The second of these issues concerns the uncertainty involved in the estimation of institutional scores from samples of students. The common convention is to indicate whether or not specified differences could be regarded as statistically significant at a specified level (typically at the five per cent level). However, the uncertainty in estimated scores (and differences

between scores) should be considered to be part of a continuum (from less to more uncertain) rather than as a simple dichotomy. This view of the uncertainty of estimates is more common in evaluation literature, where the focus is on the best possible decision, rather than research literature, where the focus is on establishing enduring generalisations. In either view it is essential to recognise that there is uncertainty in any estimate and to provide information about the degree of uncertainty.

This chapter makes use of particular approaches to handling these two sets of issues. There are other views of how these matters should be handled. Marsh and colleagues (in press) have made an important contribution to the debate about the use of the PREQ in an article and several presentations. Appendix B presents an outline of where those approaches differ from the procedures used as the basis for the analyses reported in this chapter. The approach adopted by Marsh and colleagues suggests greater uncertainty in institutional scores than reported in this chapter and therefore argues that there are fewer significant institutional differences.

---

## 5.1 Overview

An exploration of patterns of institutional results from the PREQ was conducted for each of the six scales and the Overall Satisfaction item. These analyses were restricted to the 26 institutions (out of 35) for which there were more than 20 respondents—so as to ensure sufficient numbers in each institution to provide stability in the institutional indicator. In practice this restriction resulted in analyses based on 91 per cent of all respondents from Australian universities.

A preliminary indication of the extent to which institutions differ with respect to these matters can be found in the percentages of the respondents who agreed with the items in each set. Two simple statistics provide indicators of the unadjusted variation among institutions: the range between the highest and lowest scores and the inter-quartile range between the first and third quartiles (i.e. the spread of the middle half of scores). On the Intellectual Climate scale average percentage agreement ranged from 37 per cent to 68 per cent and the inter-quartile range was 13 percentage points. For Infrastructure mean percentage agreement scores ranged from 60 to 75 per cent with an inter-quartile range of 10 percentage points. On the Supervision scale the inter-quartile range was 9 percentage points with a full range from 58 to 85 per cent. On other scales the spread was less. Inter-quartile ranges were 8 percentage points in the case of Thesis Examination (the full range was from 65 to 87 per cent), 6 points for Goals and Expectations (the full range was from 81 to 96 per cent), and 5 points for Skill Development (the

full range was from 82 to 95 per cent). For the Overall Satisfaction item there was a range from 71 to 89 per cent with an inter-quartile range of 7 percentage points.

These data serve to illustrate the magnitude of the spread of unadjusted institution percentage agreement scores. More precise measures were used in the systematic analysis of differences between institutions. Those measures involved scale scores because they were more reliable, especially where the scales contained few items, and that was considered important for the multivariate computations. However, percentage agreement (both unadjusted and adjusted) is shown in the representations<sup>9</sup>.

---

## 5.2 Procedures

In keeping with the recommendations of the report of the pilot project (ACER, 1999) and commentary on the use of the PREQ the procedure involved making allowance for differences in the composition of institutions and taking into account the spread of scores within institutions. Specifically the procedure involved three stages<sup>10</sup>:

- calculating residual scores for each respondent so as to allow for other influences,
- aggregating those residuals to institution level, and
- comparing the adjusted institutional means with a reference point.

### 5.2.1 Allowing for other influences by estimating residual scores

The analyses in this chapter took account of the influence of differences in student characteristics (level of programme, sex, age and non-English speaking background) and field of study. Field of study was represented as eight broad fields of study. Because there were very few respondents from either veterinary science or law these were grouped with science and

---

9 In fact there was a high correlation between the mean scale scores for institutions and the mean percentage agreement scores. Among the 26 institutions the median correlation coefficient was 0.90. Values of the correlation coefficient ranged from a high of 0.98 for the Supervision scale to a low of 0.79 for the Goals and Expectations scales. The other correlation coefficients were in the range from 0.87 to 0.94. For the item concerned with Overall Satisfaction the correspondence was much lower: 0.62.

10 The contribution of Dr Gerald Elsworth to the ideas used in this procedure is gratefully acknowledged.

humanities, arts and social sciences respectively. Regression equations for scale scores on programme level (doctor of philosophy or research masters), field of study, sex, age at course completion (years), non-English speaking background (NESB or ESB) and location were computed. Field of study was included as a set of dummy variables with humanities, arts and social sciences being the reference (or excluded) category. It was observed that these characteristics did not contribute greatly to the scale scores. Residuals were the difference between the estimate from the regression equation and the recorded score.

### 5.2.2 Generating institutional measures

Residuals were aggregated to institutional level as the means and the within-institution standard deviation. The within-institution standard deviation was used in conjunction with the number of respondents for the institution. The standard error for an institution depends upon the dispersion (or variability) of the scores within the institution and the number of respondents to the PREQ from the institution<sup>11</sup>.

### 5.2.3 Comparing institutional means

From the data that were generated it is possible to compare each pair of means. A difference between means was considered statistically significant if the distance between them was greater than the sum of 1.39 times the standard error of each. In presenting results institutional means and the error bars (being 1.39 times the standard error) are shown for each of the PREQ scales. Where these error bars do not overlap the difference could be considered to be statistically significant at the 5 per cent level. This provides a guide to the certainty that the difference is a real difference. However, it does not mean that instances where there is slight overlap should be dismissed as of no consequence or that where there is no overlap one can be completely certain that the difference is real. In terms of implications for policy and practice, there is no reason to assume that differences that just fail to reach significance should be ignored and those that do reach significance should prompt an automatic response.

Differences that are significant at the 5 per cent level are highlighted for guidance rather than being taken as definitive. One reason for this caution is that the data do not constitute a properly constructed sample but respondents from a population. In these circumstances, and more generally, replication will provide an additional guide as to whether observed differences represent

---

<sup>11</sup> The finite population correction was not applied although there could be an arguable case for this. The effect would be to reduce the standard errors somewhat.

underlying real differences or have arisen by chance or through bias. Of course it goes without saying that the significance level reflects the probability that the difference arose by chance. It does not necessarily reflect the magnitude of the difference. A large difference between two means based on small numbers of respondents from institutions with a wide internal variability might not be significant at the 5 per cent level. In contrast a smaller difference between two means with less internal variability, large numbers of respondents and a high response rate might be significant.

In these data there is a very large number of possible comparisons of institutional means. An alternative to comparing each pair of means is to make comparisons with the national mean. Although this provides a common point of comparison that could be estimated with some precision (because it is based on all the data), in this chapter a reference point that could be taken to represent 'good practice' was used. To operationalise this concept the institution with the second, third or fourth highest mean residual was selected as the reference point. This means that the point of comparison is not an outlier but is around the 80th or 90th percentile of residual scores. Within that range the institution chosen was one with sufficient numbers of respondents to provide a stable reference.

---

## 5.3 Results

Figures 10 through 15 show patterns for each of the six PREQ scales. For each scale three panels are shown.

- The first panel shows mean residuals for institutions and standard errors.
- The second panel shows the mean residuals and the range above and below the mean that corresponds to 1.39 times the standard error. The plot also shows the boundaries of the region plus or minus 1.39 times the standard error for a reference institution.
- The third panel shows two other indicators: the mean percentage agreement and the residual for mean percentage agreement. For these indicators the values are shown as differences from the value for the reference institution.

Figure 10: PREQ supervision scores for institutions

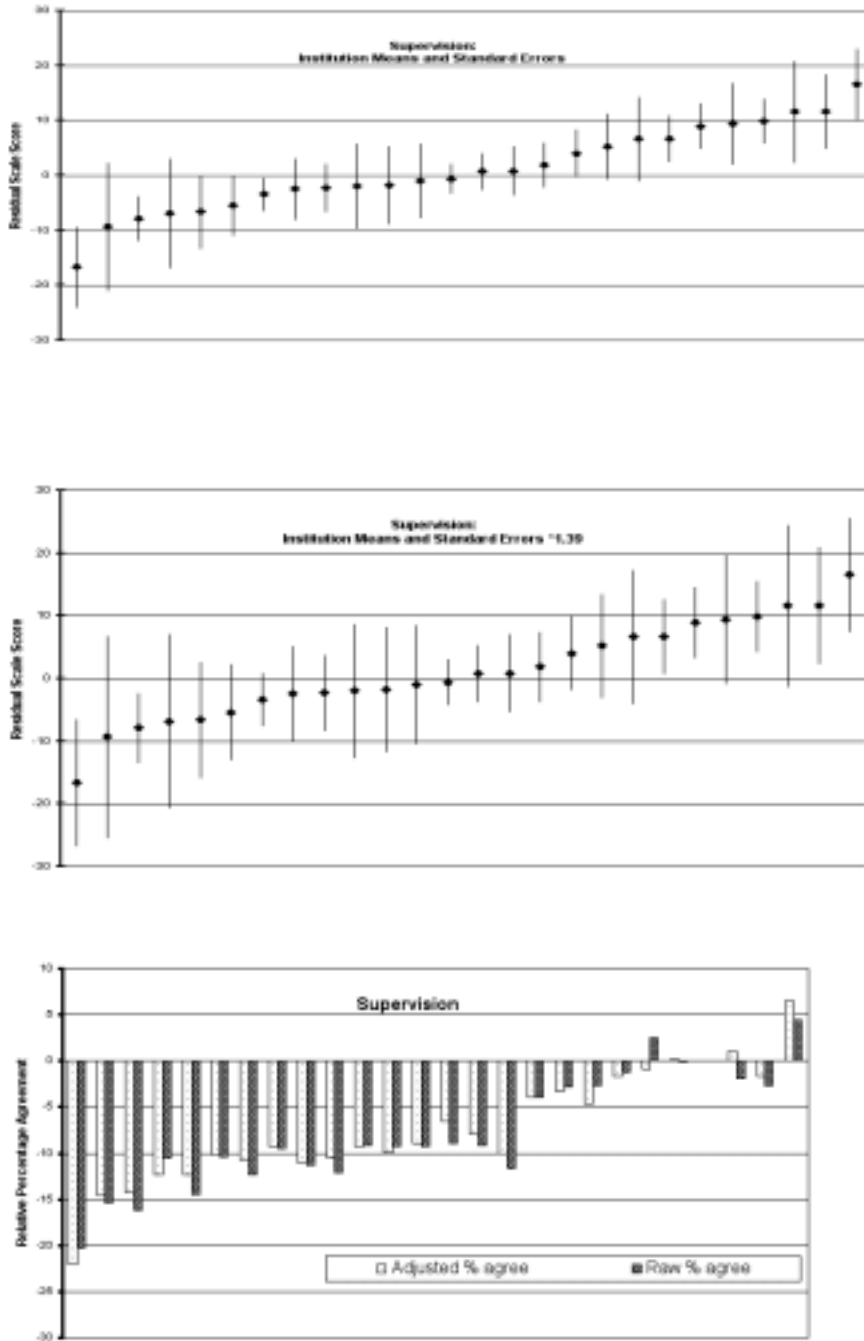


Figure 11: PREQ skill development scores for institutions

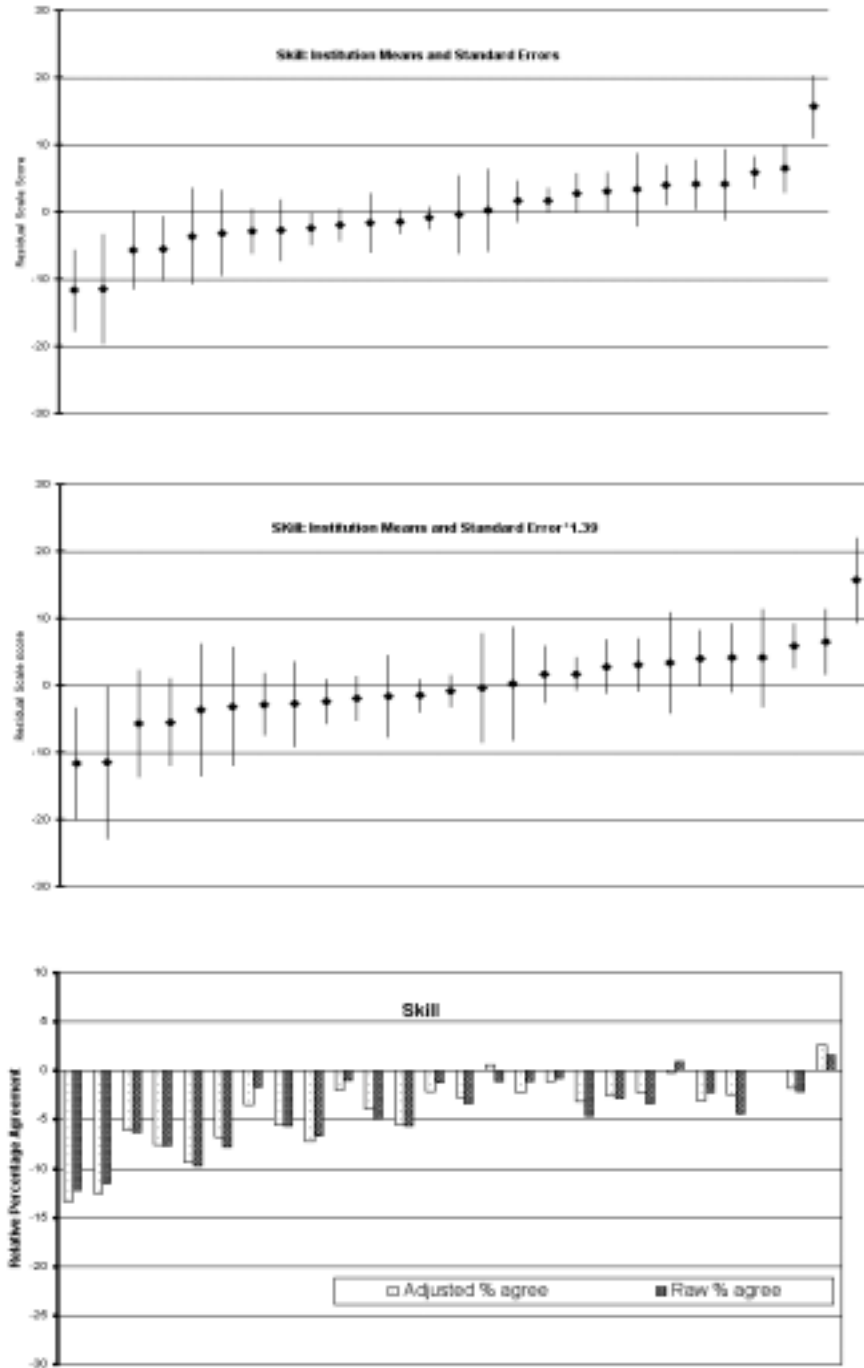


Figure 12: PREQ intellectual climate scores for institutions

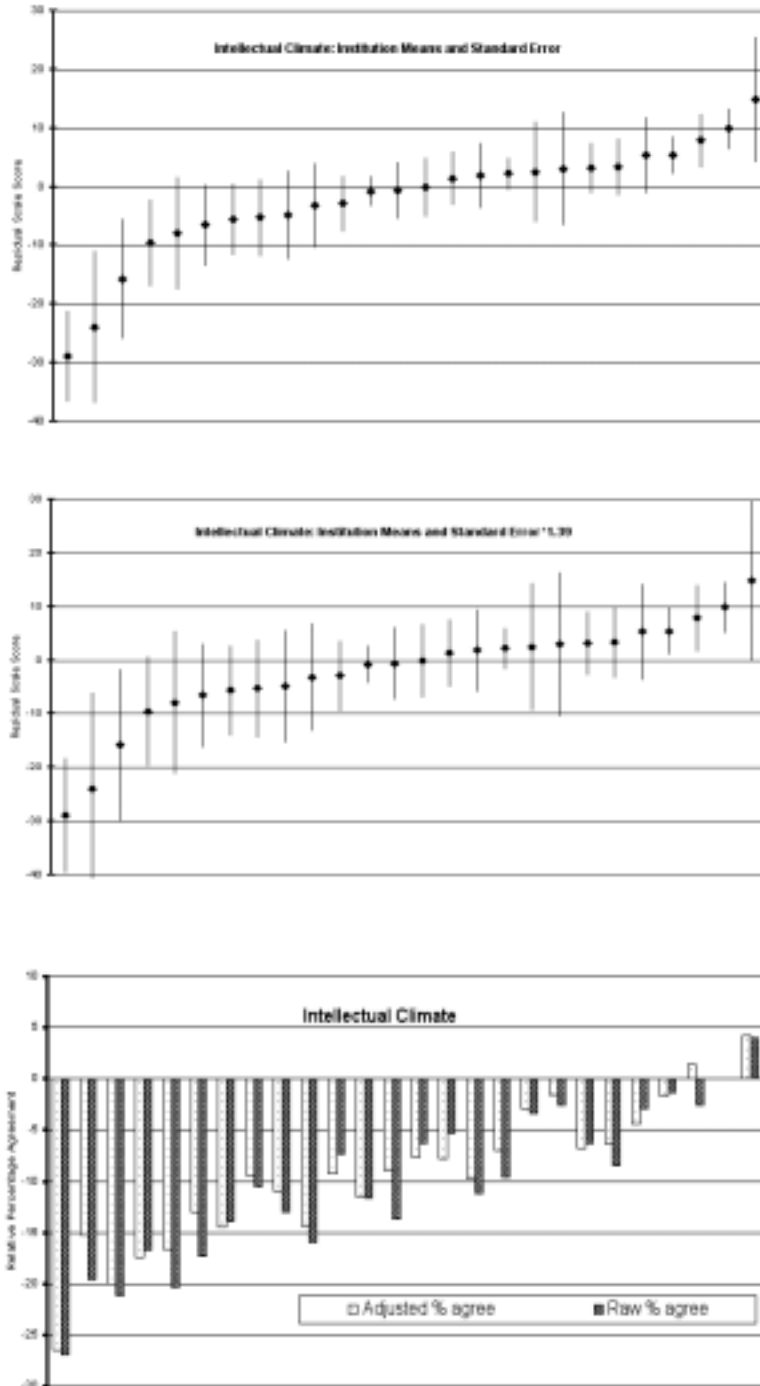


Figure 13: PREQ infrastructure scores for institutions

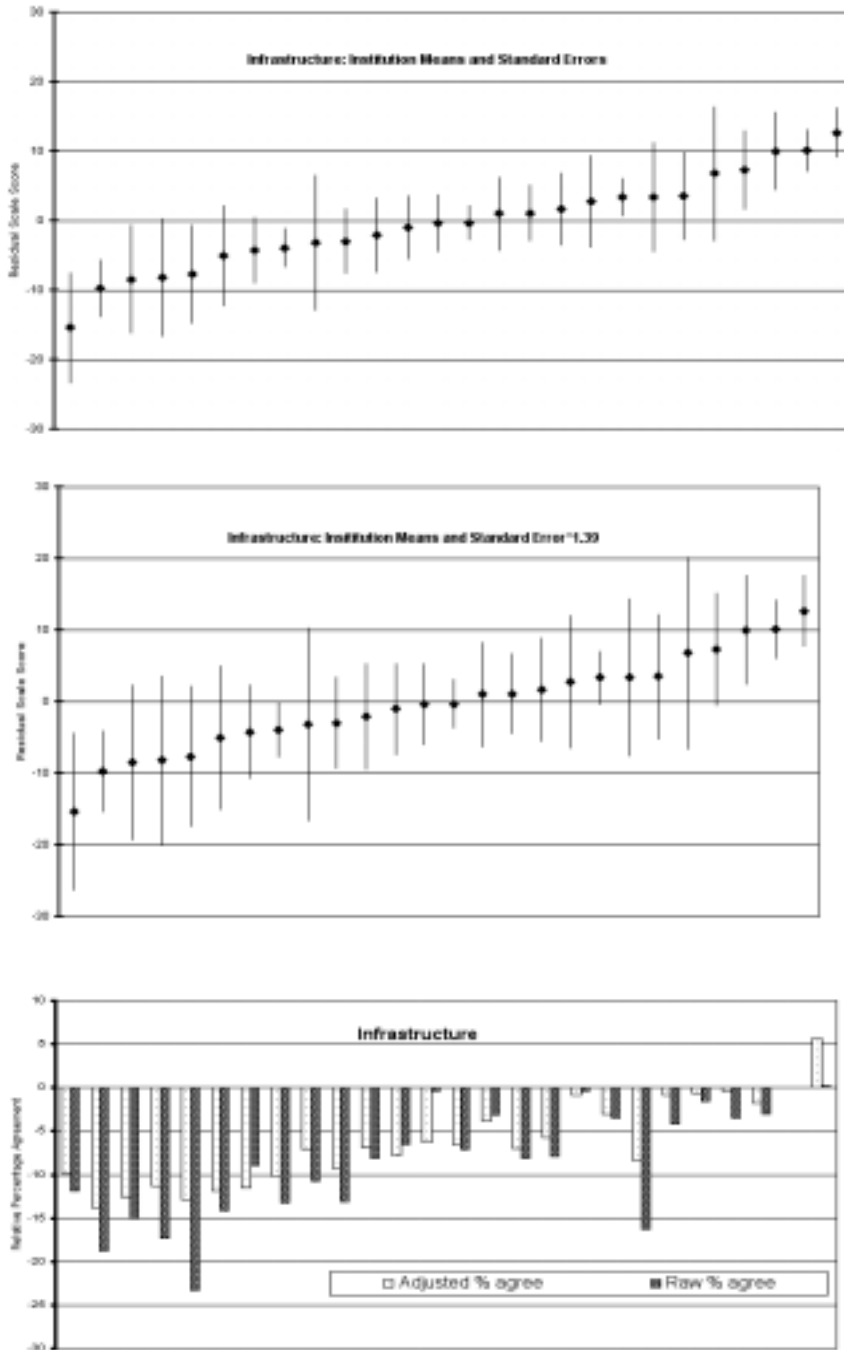


Figure 14: PREQ thesis examination scores for institutions

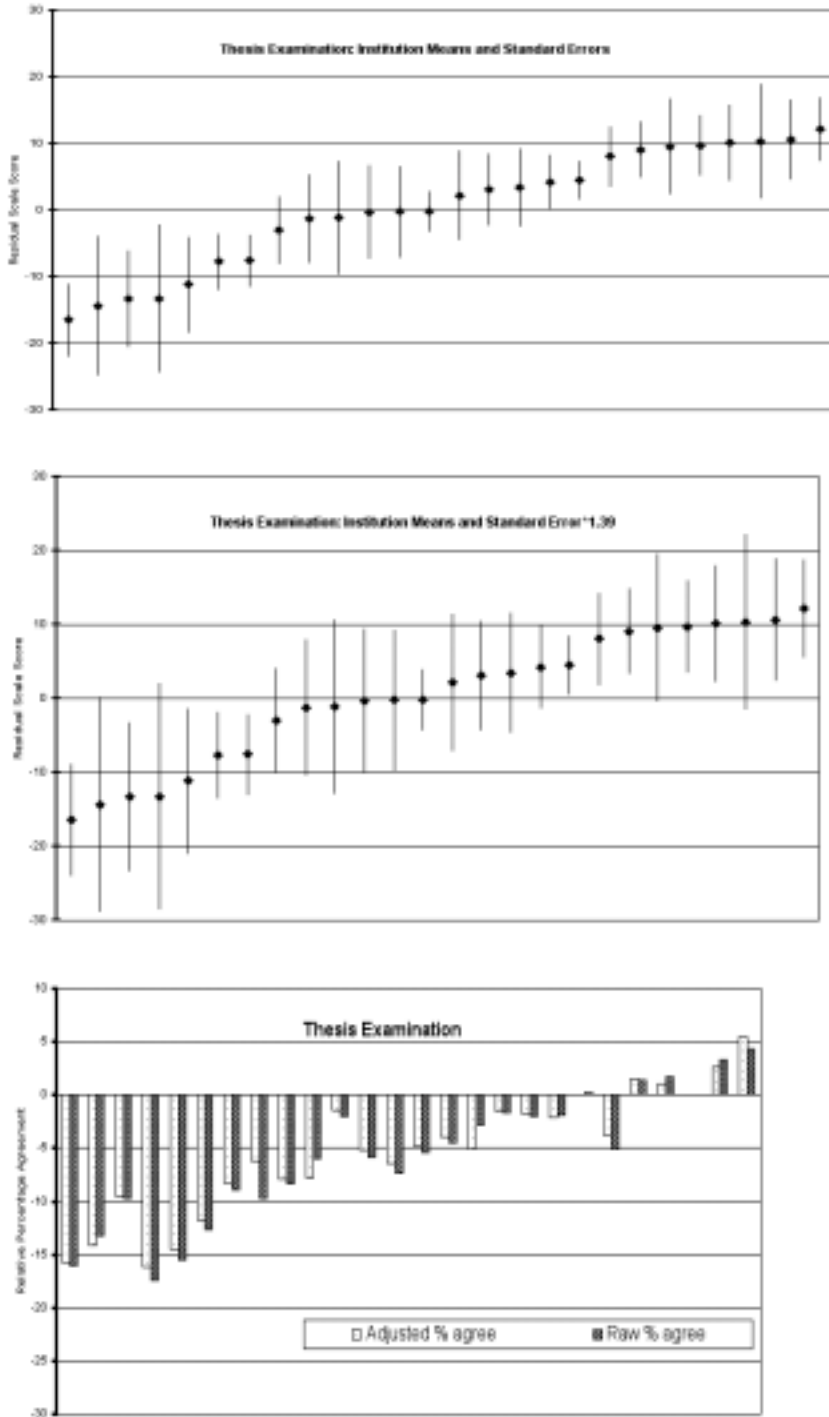


Figure 15: PREQ goals and expectations scores for institutions

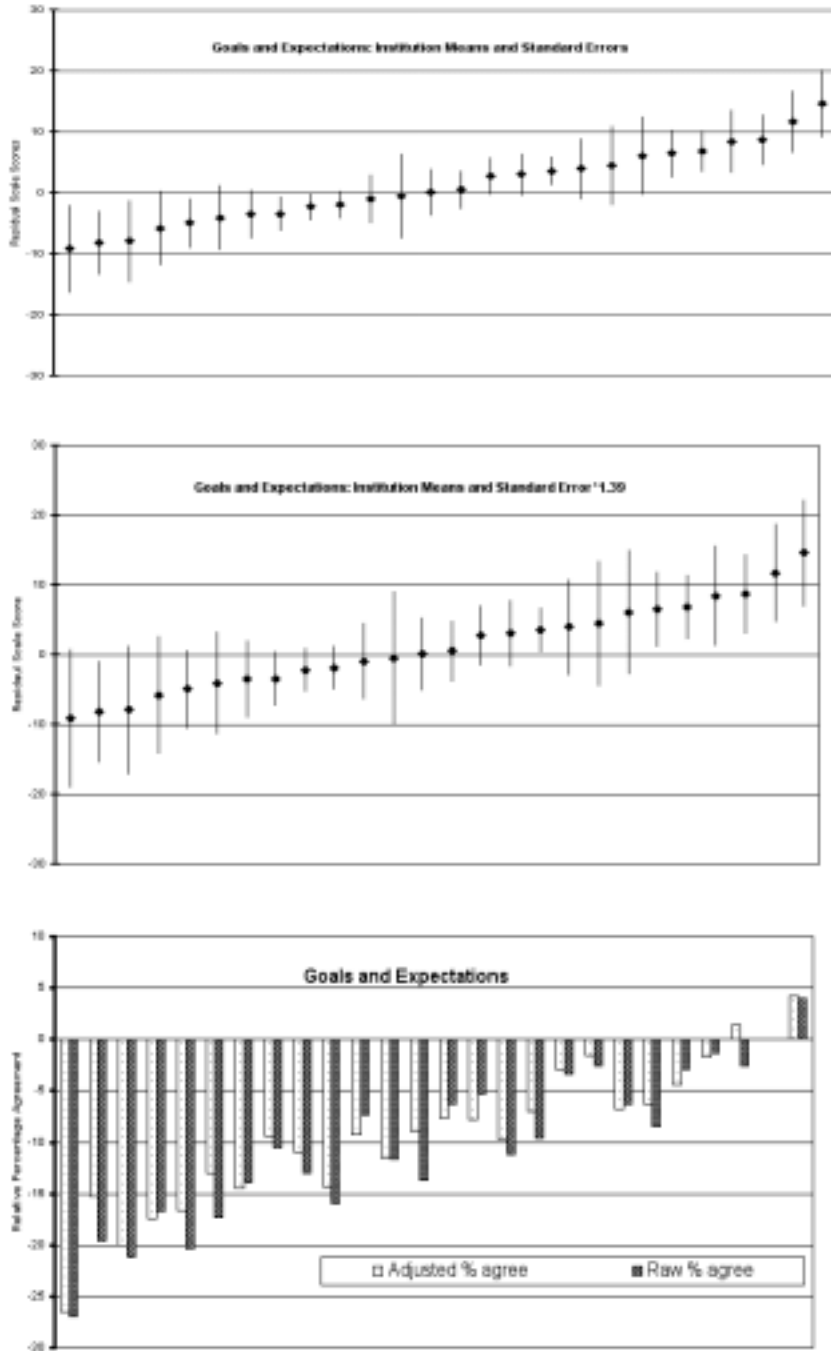


Figure 16: PREQ overall satisfaction scores for institutions

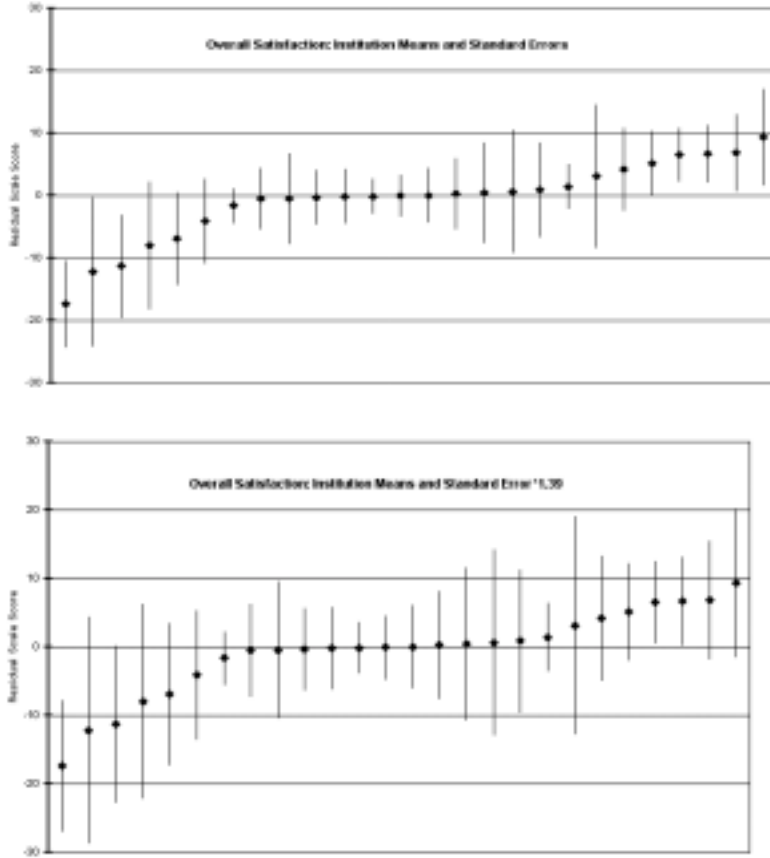


Figure 16 shows the pattern of Overall Satisfaction scores across institutions. There are only two institutions where the error bars do not enter the boundaries of uncertainty for the reference institution. On the basis of these data it would appear that the Overall Satisfaction item does not discriminate between institutions to the same extent as several of the scales. However, even for this item there are instances where differences between pairs of institutions are evident.

---

## 5.4 Summary

The analysis of PREQ indicators for 26 Australian institutions in which there were more than 20 respondents suggests that there may at least be a case for examining institutional differences further.

The differences in some of these indicators appear to be of moderate size. Of course these differences could arise from factors other than the postgraduate research experience. Information about the range of percentage agreement scores provides an indication of the extent of institutional differences unadjusted for any differences in the composition of the research graduates (on such factors as the mix of fields of study, the proportions of masters and doctoral graduates, and background characteristics). In addition they do not show the extent of within-institution variation.

Analytic techniques were applied to adjust for differences in field of study and level of course as well as background factors. Estimates were made of within-institution variation so as to inform inferences about whether observed differences among institutions should be considered to represent real differences or the effects of chance. As a result of these analyses it appeared that there was greater institutional variation on the Intellectual Climate scale than on the others. On most scales there were a number of institutions that could not be distinguished on the basis of this one year of data. However, it was possible to observe on each scale that there were institutions that differed significantly from a reference point selected to represent good practice. Although these data would not support the compilation of an ordered list of institutions they do suggest the possibility that the results could be used to establish where there is a discrepancy between an institution and good practice. That application would be stronger if it were based on replicated results from more than one survey year.

## 6. Conclusion

The data for this report were derived from the first full national survey of graduates of research-based postgraduate degrees. It follows a development and trial process that was extensive (ACER, 1999) and a critique of its application (Marsh, Rowe & Martin, in press).

Graduates who completed the doctor of philosophy degree or a master's degree by research in 1998 constituted the survey population. More than 2000 graduates from 35 Australian universities and one New Zealand university returned surveys. The present report focuses on the graduates from Australian universities, and where institutional patterns are considered, on the 26 Australian universities for which there more than 20 respondents. Although the overall response to the survey was approximately 50 per cent it did appear that the respondents to the survey matched the population of course completions reasonably well on a number of course and background characteristics.

Among the issues involved in the PREQ it seems appropriate to separate those concerned with the questionnaire itself from those concerned with its application. Issues of application would include whether scores on the PREQ are related to characteristics of the respondents and the courses they completed and whether those scores can be used to discriminate between universities. In addition there is a set of issues concerned with the extent to which inferences can be drawn from PREQ data about institutional learning environments.

The information in this report supports the view that the PREQ in its present form is a sound instrument for gathering the views of graduates as individuals from research based higher degrees about the domains that it samples. The internal structure of the instrument that emerged from the developmental trials was confirmed in the analyses of these data and the scales seem to have a high level of internal consistency. It could be argued that the focus should be more on the four scales that contain five or six items on the grounds of robustness but that is a matter for the emphases that is placed on different domains in reporting. On the basis of this, and given that issues of supervision, skill development, intellectual climate and infrastructure are central components of the quality of postgraduate learning environments, the PREQ seems an appropriate instrument for gathering and reporting data at a national level.

A central feature of the conception of the PREQ is its multi-factorial nature. There is evidence of a second order construct underlying its scales but in this report the separate scales were maintained through the analyses at both individual and institutional level.

Differences in PREQ responses associated with differences in the background characteristics of respondents were neither extensive nor large. The biggest differences concerned Infrastructure although it should be noted that some of these differences reflected concomitant differences in field of study.

Differences associated with graduate background could arise from differences in the types of programme in which students are enrolled, differences in perceptions of the same situation or differences that arise from variations in programmes or treatment within programmes.

Differences between doctoral and masters degree graduates and between fields of study were observed. The differences between doctoral and masters graduates were in the expected directions for each scale. Differences among fields of study are more difficult to interpret. Moderate differences between fields of study were evident in Overall Satisfaction, Infrastructure, Supervision and Intellectual Climate. The differences were not the same across the domains. Graduates from science related fields of study tended to have higher scores on the Infrastructure scale. That result, and the low rating accorded to the Infrastructure domain by postgraduates in the arts, humanities and social sciences, is consistent with what one might expect. The field of the arts, humanities and social sciences was also relatively low in terms of Intellectual Climate and Overall Satisfaction. Education postgraduates also rate their experience poorly in terms of Infrastructure support but highly in terms of Supervision and Overall Satisfaction and reasonably in terms of Intellectual Climate.

The analysis of institutional patterns suggested that it was possible to identify on each scale a number of institutions that differed from a reference institution selected to represent good practice. The extent of differences was greater on some scales than others with greater institutional variation on the Intellectual Climate scale than on the others. This is not inconsistent with there being a range of institutions that could not be distinguished from each other. Although these data would not support the compilation of an ordered list of institutions they do suggest the possibility that the results could be used to establish discrepancies from good practice. Even then the appropriate use of such information would be to stimulate further inquiry using other methods. More importantly one would wish to ensure that a discrepancy was replicated in results from more than one survey year before deriving inferences about institutional learning environments.

Use of accumulated data over several years would be valuable for establishing replicated patterns that provide greater confidence in the inferences that are drawn from data such as these. Such accumulations are important in a situation where the total numbers involved are small and the numbers in particular categories are smaller still.

## References

Australian Council for Educational Research (ACER) (1999). *Evaluation and validation of the trial Postgraduate Research Experience Questionnaires*. Canberra: DETYA.

Elsworth, G. (1999). *The development and properties of the postgraduate research experience questionnaire*. Paper presented to the seminar on the PREQ. Melbourne: Graduate Careers Council of Australia, December.

Marsh, H., Rowe, K. & Martin, A. (in press). PhD students' evaluations of research supervision: Issues, complexities and challenges in a nationwide Australian experiment in benchmarking universities. *Journal of Higher Education*.

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, 21, 173–184.

Wilson, K., Lizzio, A. & Ramsden, P. (1997). The development, validation and application of Course Experience Questionnaire. *Studies in Higher Education*, 22 (1), 33–53.

# **Appendix A: The Postgraduate Research Experience Questionnaire 1999**

The purpose of these questions is to collect information relating to graduates' perceptions of their higher education experience. Please complete the following items on the basis of your most recently completed degree (the one you listed in Question 1 on the front of this survey form).

In some instances, higher degree research graduates will have had more than one supervisor and/or have studied in more than one department and/or faculty. If this is so for you, we are asking that your responses regarding supervision relate to your most recent supervision whether by one or more supervisors. Please interpret 'thesis' and other research-related terms in the context of your own field of study.

The 1999 Postgraduate Research Experience Questionnaire

To answer, please circle the number beside each statement that most accurately reflects your view. You may choose from a scale where '1' means that you strongly DISAGREE with the statement and '5' means that you strongly AGREE. Circle '9' if the statement does not apply to your experience.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree	Does not apply
1. Supervision was available when I needed it	1	2	3	4	5	9
2. The thesis examination process was fair	1	2	3	4	5	9
3. I had access to a suitable working space	1	2	3	4	5	9
4. I developed an understanding of the standard of work expected	1	2	3	4	5	9
5. The department provided opportunities for social contact with other postgraduate students	1	2	3	4	5	9
6. My research further developed my problem-solving skills	1	2	3	4	5	9
7. My supervisor/s made a real effort to understand difficulties I faced	1	2	3	4	5	9
8. I had good access to the technical support I needed	1	2	3	4	5	9
9. I was integrated into the department's community	1	2	3	4	5	9
10. I learned to develop my ideas and present them in my written work	1	2	3	4	5	9
11. I understood the required standard for the thesis	1	2	3	4	5	9
12. I was able to organise good access to necessary equipment	1	2	3	4	5	9
13. My supervisor/s provided additional information relevant to my topic	1	2	3	4	5	9
14. My research sharpened my analytical skills	1	2	3	4	5	9
15. I was satisfied with the thesis examination process	1	2	3	4	5	9
16. The department provided opportunities for me to become involved in the broader research culture	1	2	3	4	5	9
17. I was given good guidance in topic selection and refinement	1	2	3	4	5	9
18. I had good access to computing facilities and services	1	2	3	4	5	9
19. I understood the requirements of thesis examination	1	2	3	4	5	9
20. Doing my research helped me to develop my ability to plan my own work	1	2	3	4	5	9
21. My supervisor/s provided helpful feedback on my progress	1	2	3	4	5	9
22. A good seminar programme for postgraduate students was provided	1	2	3	4	5	9
23. The research ambience in the department or faculty stimulated my work	1	2	3	4	5	9
24. I received good guidance in my literature search	1	2	3	4	5	9
25. The examination of my thesis was completed in a reasonable time	1	2	3	4	5	9
26. As a result of my research, I feel confident about tackling unfamiliar problems	1	2	3	4	5	9
27. There was appropriate financial support for research activities	1	2	3	4	5	9
28. Overall, I was satisfied with the quality of my higher degree research experience	1	2	3	4	5	9

What were the best aspects of the degree? Please write below:

---



---



---

What aspects are most in need of improvement? Please write below:

---



---



---

Office use only: optional course/faculty variable ( ) ( ) ( ) ( ) ( ) ( ) ( ) 50-57

## Appendix B:

# Issues in the analysis of institutional means

In the analyses presented in this report, and especially in the analyses concerned with institutional differences presented in Chapter 5, choices regarding analytic procedures were made. Since the procedures adopted have been criticised (Marsh et al, in press), this Appendix presents some key elements of the critique and provides a little more detail regarding the basis for the approach than was adopted in the present report. An integrated account of the methodology advocated in the critique can be found in an article by those authors (Marsh et al, in press). The critique is based on data from the trial of the PREQ but subsequent presentations have indicated that analyses based on the present national survey do not differ in any substantial way from those of the trial data.

### The multilevel nature of PREQ data

PREQ data are obtained from the responses of individual graduates and those data are used to generate indicators at institutional level. The structure of the data is such that students are nested within universities and this needs to be recognised in the analytic procedures. This is common in educational research and poses challenges for the analysis of data. Multilevel analysis allows simultaneous consideration of multiple units within the same analysis. Marsh and colleagues (in press) use the widely known MLwiN and LISREL packages for this analysis and to test the effects of other variables on scores. In the simplest models the graduate would be level 1 and the university would be level 2. They point out that 'the selection of the appropriate units of analysis is not straight forward' (especially for a variable such as field of study). They fitted a series of two and three level models to the PREQ data. Using these multilevel analyses, they reported that there were no significant differences between universities on the PREQ responses. Hence, they concluded that PREQ responses are unreliable for purposes of benchmarking universities.

This approach to the analysis of the PREQ data appears to implicitly assume that those data are based on a random sample of graduates from each university and a random sample of universities. A common situation in which these approaches are applied involves a two-stage (or possibly three-stage) sample where schools are sampled from a population of schools and students are sampled within schools. The analyses presented in Chapter 5 do not

consider the PREQ data to be based on a sample of universities (although it can be argued to involve a sample of graduates from those universities). The survey was based on Australian universities where each individual institution is better regarded as constituting a 'fixed treatment' in its own right rather than one randomly sampled from a large population of 'treatments' (institutions). It would be possible to envisage these data as representing a sample of universities over time or of a meta-population of possible Australian universities. However, the view adopted in this report was that there was no intention to generalise beyond this year and this particular group of institutions.

### Confidence intervals for the comparison of institutional means

Comparing institutional means involves estimating the errors associated with those means. The error associated with an institutional mean depends on the number of respondents for the institution (larger numbers result in smaller errors) and the dispersion (or variability) of responses within the institution (greater internal variability results in larger errors). The precision in an institutional mean is reflected in the standard error for that mean. In the 2000 PREQ symposium it was argued that two means should be regarded as significantly different from each other if the intervals based on 1.96 times the standard errors did not overlap. The critique argues that there are many pair-wise comparisons that are possible and the approach needs to reflect that some differences could arise by chance. The approach adopted in the present report was that two means could be regarded as significantly different from each other if the intervals based on 1.39 times the standard error do not overlap. This is because one is considering pair-wise (rather than multiple) comparisons. The issue is clearly of importance to the inferences that can be drawn from these PREQ data.

It can also be noted that the analyses in Chapter 5 have not applied a finite population correction to the error estimates. If a finite population correction were to be applied it would result in the estimated precision of the estimates of institutional means being somewhat better. The finite population correction recognises that the closer the sample is to the full population the smaller is the error (and if it is identical with the full population the sampling error becomes zero). In many surveys the sample is much smaller than the population and the finite population correction is not an issue. In the 1999 PREQ the average response rate for universities was 50 per cent. Applying the finite population correction to an institution with a 50 per cent response rate would reduce the standard errors, and hence the confidence intervals, to 0.7 of those estimated in Chapter 5. In institutions where the response rate was 75 per cent the standard error would be just half the value estimated and

where the response rate was 25 per cent the effect would be to reduce the error to 0.87 times that estimated.

### Field of study

If there were a larger data set on which to base analyses it might be possible to analyse institutional differences within each field of study. However, with a limited number of respondents in many fields of study it is necessary to make a statistical allowance for the influence of field of study. In the analyses presented by Marsh and colleagues (2000) this is accomplished by including dummy variables representing field of study in the analytical models. In the present Chapter 5 regression methods were used to estimate residuals for individual respondents that were then aggregated to institutional level to accomplish the same purpose. There is some ambiguity as to how field of study should be conceptualised for this purpose. In practice the differences between fields of study appear to be smaller than has been observed on the CEQ.