



Research Report

Unlocking the impact of tertiary teachers' perceptions of student evaluations of teaching

Sarah J. Stein, Dorothy Spiller, Stuart Terry,
Trudy Harris, Lynley Deaker and Jo Kennedy

Authors

Sarah Stein, Dorothy Spiller, Stuart Terry, Trudy Harris, Lynley Deaker, Jo Kennedy

Publishers

Ako Aotearoa National Centre for Tertiary Teaching Excellence
PO Box 756
Wellington 6140

This project was supported through the Ako Aotearoa National Project Fund 2009, in the Research and Implementation Projects funding stream. The Guide *Using student evaluations to enhance teaching practice: Closing the loop* (Stein et al., 2012) complements this Research Report.

Published

September 2012

Key Words

student feedback; formal institutional evaluations/appraisals systems; teacher perceptions; worth of student evaluations; perceptions of limitations; engagement with and use of student feedback; professional development; feedback to students aligning policy, perceptions and practice.

<http://ako.aotearoa.ac.nz/student-evaluations>

ISBN: 978-1-927202-15-9



This work is published under the [Creative Commons 3.0 New Zealand Attribution Non-commercial Share Alike Licence \(BY-NC-SA\)](https://creativecommons.org/licenses/by-nc-sa/3.0/). Under this licence you are free to copy, distribute, display and perform the work as well as to remix, tweak, and build upon this work non-commercially, as long as you credit the author/s and license your new creations under the identical terms.

Key Words

student feedback; formal institutional evaluations/appraisals systems; teacher perceptions; worth of student evaluations; perceptions of limitations; engagement with and use of student feedback; professional development; feedback to students aligning policy, perceptions and practice.

Executive Summary

Introduction

This study explored the extent to which teacher thinking and practices at all stages of the teaching and learning cycle are informed by learner feedback gathered through the institutional student evaluation process. This process includes the standard student evaluation/appraisal of teaching and course questionnaires and the associated policies, systems and procedures administered centrally by institutions. This investigation of the influence of student evaluations/appraisals on teacher professional development and student learning aimed to contribute directly to Ako Aotearoa's National Project Fund 2009 priority area of Strategic Professional Development Initiatives. A core goal of this project was to provide insights to assist tertiary institutions and individuals in order to strengthen the relationship between student feedback, teaching development and the quality of the student learning experience.

Origins of the Research and Research Goals

The researchers' experience of working with staff in their own institutions prompted a number of questions around teacher views about, and use of, student evaluations/appraisals. These questions included the potential impact of teacher perceptions of the institutional use of evaluations/appraisals data for quality monitoring purposes and for promotion considerations. Other sentiments that the researchers wanted to probe were academic suspicion of students' capacity to make judgements about their learning and scepticism about the quality of evaluation instruments. Additionally, the research team was motivated by a concern that there might be major shortcomings in teacher engagement with student feedback from formal evaluation processes and that student evaluation/appraisal data were not broadly used to inform teaching development and the quality of the student learning experience.

Finally, the researchers were interested to find out to what extent students were included in ongoing dialogue around evaluation/appraisal.

These initial questions and concerns provided the genesis for the study and evolved into the broad research question:

How do current evaluation processes and practices influence teachers' thinking and behaviours in relation to student learning at all stages of the teaching and learning cycle?

and the contributing sub-questions:

- 1. What are the perceptions that tertiary teachers hold about student evaluation/appraisal?*
- 2. What factors (causes, influences) affect these views?*
- 3. How do tertiary teachers engage with student evaluations/appraisals?*

Design and Methods

In order to investigate teacher perceptions and engagement and use of student evaluation feedback, the study explored tertiary teachers' perceptions about evaluations/appraisals in three institutions, the University of Otago (OU), The University of Waikato (WU) and the Otago Polytechnic (OP). The researchers were academic developers and appraisals personnel from these three institutions.

An interpretivist research approach framed the study that drew on a combination of quantitative and qualitative data, including a questionnaire and interview. The research questions, the questionnaire design and the interview questions were informed and sharpened by a literature review, as well as an environmental scan of evaluation policies and practices made public by New Zealand polytechnics and universities through their websites. The literature review focussed, in particular, on questions of teacher perceptions of, and engagement with, student evaluations and provided a benchmark against which the researchers evaluated the research findings that emerged as the study progressed. The broad trends in the literature indicated a reasonably positive disposition among academics towards undertaking formal evaluations/appraisals. They also reinforced the caveats that had been reported in the institutions participating in this study and suggested that engagement with,

and use of, student evaluation data was very limited, lagging well behind the apparently positive perceptions.

An online questionnaire was run with academic staff over a period of three weeks. In all, 2426 teaching staff from across the three institutions were invited to respond to the questionnaire and 1065 responses were received (44 per cent). Statistical analysis was carried out on the responses to the Likert-style questions of the questionnaire and thematic analysis was used to investigate responses to the open comment questions.

Semi-structured interviews were subsequently conducted using questions that were developed on the basis of the themes that emerged from the questionnaire findings. Sixty interviewees were selected from volunteers – 20 from each institution – to provide as broad a sample as possible in terms of academic discipline, career stage and seniority. The interviews enabled the implications of the themes that became evident from the questionnaire data to be explored in more detail on a one-to-one basis. As with the questionnaire comment data, thematic analysis was used to draw out key themes from the interviews.

Through an analysis of the findings from all across all data sets, in the light of the literature, conclusions could be made about perceptions held by staff between and across all participating institutions.

Findings

Sub-questions 1 and 2: Teacher perceptions of evaluations/appraisals and reasons for these views

In line with the literature review, and somewhat counter to the researchers' initial hunch, there was widespread recognition in the questionnaire responses that participants thought that it was worthwhile collecting evaluation/appraisals data (73 per cent). The most commonly identified beneficial reasons were to inform teacher and/or course development (19 per cent) and identify students' learning needs (19 per cent). Of the limitations identified, the most common included: shortcomings in the current evaluations/appraisals system (14 per cent); quality of student responses (10 per cent); and the use of the same instrument for quality and development purposes (seven per cent). The relatively small numbers of respondents who expressed a concern about the dual purposes of evaluations/appraisals was a surprise to the

researchers, who had received a different subjective impression from their contact with academic staff.

The 60 interviews provided further illumination about academics' perceptions of evaluations/appraisals, thereby reinforcing the findings of the questionnaire and suggesting some other aspects that needed further exploration. A strong reservation voiced in the interviews at the two universities (nine at OU; 11 at WU) concerned the quality of student feedback, with a smaller number questioning the quality of student judgements at OP (six). At the universities there was also concern about limitations in the institutions' evaluation/appraisal survey instruments (eight at OU; eight at WU), but this was not a noticeable concern in the OP interviews. Another interview finding that was raised by OU interviewees was the potential manipulation of the evaluations/appraisals process by academics (eight). Other problems mentioned were timing of evaluations/appraisals, and the associated unease about institutional use of the data (eight at OU; three at WU; and two at OP), including a concern about institutional reliance on one evaluation source to base assumptions and decisions. A contextual difference that emerged in interviews with OP interviewees was that small classes and close contact in skills-based teaching meant teachers had many opportunities to gather informal feedback on their students' learning and consequently formal evaluation/appraisal data were seen as less significant.

The interviews purposefully explored two additional areas, namely, teaching philosophy and the personal response that may be generated by the evaluations/appraisals process. With regard to teaching philosophy, at the universities there was some correlation between academics whose expressions about their teaching conceptions were reflective and student-focussed, and their positive views about students' capacity to make judgements in evaluations/appraisals. This correlation was not the case for OP interviewees, 13 of whom, in spite of their focus on student learning needs, expressed ambivalence about students' capacity to judge the quality of the teaching and learning. Possibly, because a student-focussed philosophy appears to be much more the norm in the polytechnic context, it is perhaps a less obvious potential indicator of teacher attitudes to evaluation/appraisal feedback. Another interesting contextual difference was the high number of interviewees from OU who prioritised discipline research in their teacher conceptions (eight) over research into teaching.

The implications of this perception and its impact on views of student evaluation/feedback merits more detailed investigation.

Finally, prompted by the original impetus for this study, the research literature and the glimpses of intense emotional language in the questionnaire comments, the interviews specifically tried to uncover the role of personal responses or emotions around the evaluations/appraisals process. The language used by interviewees indicated that for some academics there is a strong affective (positive and negative) element in the evaluation/appraisals process (10 at OU; seven at WU; and six at OP). Because of the seemingly important role of emotional and personal responses to the evaluation/appraisal process revealed by the study, this aspect could be another focus for future research.

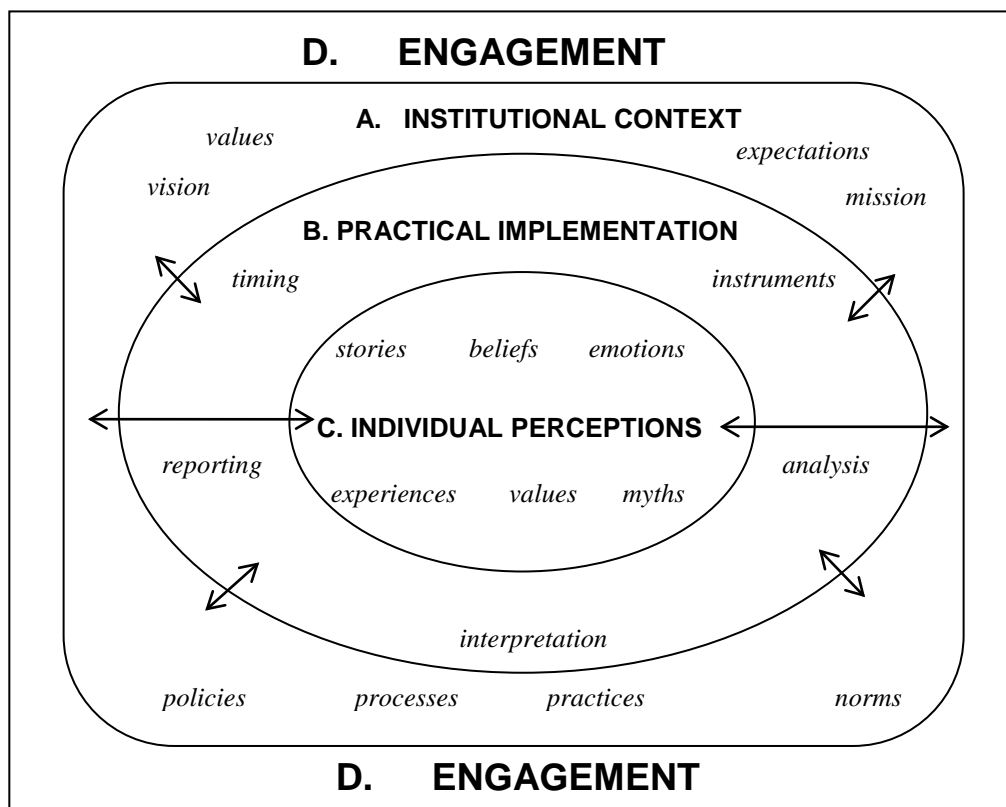
Sub-question 3: How do tertiary teachers engage with student evaluations/ appraisals?

The questionnaire indicated that, out of a range of reasons for using student evaluations, academics at each of the two universities ranked getting feedback on students' learning experiences as their number one reason. By contrast, respondents from OP ranked 'because it is required' as highest. The majority of participants across all institutions indicated that reading the open questions/comments was an activity they undertook when they received results from their evaluation/appraisal surveys (95 per cent). High numbers of participants said they spend time going over the data (87 per cent), looking for feedback on teaching (77 per cent) and comparing data with previous results (77 per cent). By contrast, numbers of respondents who said that they sought assistance with interpreting results were relatively low (12 per cent). In particular, talking or working with colleagues or academic developers (47 per cent) and feeding their evaluation trends and responses back to students (16 per cent) were not activities in which respondents said they engaged to high degrees. The culture of a private, isolated engagement with evaluations/appraisals data was particularly evident in the universities, with evidence of more engagement at OP. At all three institutions, the reason most commonly cited for not feeding back to students was that evaluations/appraisals tended to be run at the end of a semester, and therefore their timing made reporting back to students impossible.

These questionnaire findings in relation to engagement and use of evaluations/appraisals data were confirmed by the interviews. The interviews did reveal that at all three institutions there

was a small number of staff who have deliberate and systematic ways of engaging with feedback, use it to inform professional development and methodically share evaluation/appraisal findings with students. However, those who claimed they followed these steps were in a minority. Timing was cited as a significant limitation, and the prevalence of an isolated, private cultural norm around evaluations/appraisals data was confirmed. Interviewees from OP reported a higher degree of engagement with student feedback than those from the universities. Even so, the number of interviewees who engaged in this way, from all three institutions, was not high. Another commonly raised concern was that centralised student evaluation/appraisal systems only offered one form of data gathering and that there was a need for multiple forms of evaluation to gain a broader view of the effectiveness of teaching and courses. In some cases, academics who regularly used other forms of evaluation ignored the feedback from the formal systems, while in other cases academics suggested that using multiple forms of evaluation helped to improve the usefulness of the formal evaluations/appraisals. Generally, the interviews confirmed that there is a gap in the quality of engagement with evaluations/appraisals data, its deliberate and systematic use for professional development and ongoing engagement with students about their feedback and how it is being valued and used.

To summarise the research findings, a conceptual framework was developed to illustrate, diagrammatically, the relationships among: *individual perceptions* (C) and the character of those perceptions; the *practical implementation* (B) factors surrounding the student evaluation instruments, policies and processes; and the *institutional context* (A) including expectations (for example, expressed through vision, mission and values, and more specifically through evaluation policies, processes and norms). These factors have a bearing on the nature of teacher *engagement* (D) in student evaluation. See below.



Recommendations

The findings of this study provided a foundation for a series of recommendations about how to move towards a shared understanding of teaching and teaching evaluation that could enhance the capacity of centralised student evaluation/appraisal systems to improve teaching and student learning environments. The research suggested the need to design and implement teaching evaluation approaches that are collaborative and organic, not solitary and isolated. Such an endeavour could be complementary to, and not undermined by, the quality purposes of student evaluation/appraisal.

The recommendations are for institutions, the units within institutions (faculties/departments/schools), academic staff developers and for individual tertiary teachers. They are organised around a series of key areas that the study has shown to be influential in determining tertiary teachers' perceptions of student evaluation/appraisal. They focus on enabling teachers and institutions to optimise their use of student feedback gained through student evaluation/appraisal and to align activities – whether related to development

of teaching or accountability – more effectively with learners’ needs. The recommendations call for a concerted effort to align: the conceptualisation of evaluations held by all stakeholders; the representation of these conceptualisations in institutional policies and guidelines; and the enactment of corresponding behaviours by the institution and teachers. The recommendations are as follows:

Recommendation 1: *That institutions ensure there is a clear alignment between their vision/policy statements concerning the auditing and developmental purposes of student evaluation/appraisal systems and their processes of implementation.*

Recommendation 2: *That institutions implement a professional development strategy that includes explicit support for the education of staff and students about the purpose of student evaluation for curriculum and teaching, and the institutional intents and purposes of its student evaluation/appraisal system.*

Recommendation 3: *That those who administer student evaluation systems recognise and acknowledge the variety of staff perceptions about student evaluation/appraisal and provide communication, support and resources that address teacher expectations and needs, without compromising institutional intents and purposes.*

Recommendation 4: *That institutions ensure expectations about teacher and student roles and responsibilities in evaluation are unambiguous, and connections among performance, evaluation and reward are clearly understood.*

Recommendation 5: *That teachers, faculties/departments/schools and institutions embed within evaluation policies and practices the notion that a ‘well-rounded’ representation of teaching and courses is more likely to be achieved by drawing on multiple forms of evaluation data.*

Recommendation 6: *That professional development and course enhancement are firmly ensconced as the foundation and foci of student evaluation processes and practices. Institutions should devise a system that clearly defines the developmental and auditing purposes of student evaluation. The system should include processes and practices that target each purpose, but that also recognise that the purposes are complementary in nature and that a level of integration is needed to provide cohesion.*

Further Research

This study provided an overview of a wide range of tertiary teachers' views about student evaluations/appraisals, and the ways that tertiary teachers say they engage with student evaluations/appraisals processes. Some assumptions held by the researchers and appearing in the literature were confirmed, while others were challenged. Inevitably, the research has opened up many further lines of inquiry that need to be explored in order to maximise the long-term benefits of this study for all concerned in tertiary teaching and learning, its evaluation and development. Some areas for further research that this study identified are:

- possible influences of disciplines and professions on teacher perceptions of student evaluation/appraisal
- personal and emotional aspects of involvement in evaluation/appraisal
- possible connections between and amongst gender, age, ethnicity and perceptions of evaluation/appraisal
- relationships between perceptions of evaluation/appraisal and career progression
- differences and similarities in perceptions about evaluation/appraisal when roles change to include more management than teaching
- similarities and differences between perceptions of those who hold teaching-only positions and those who hold management or administration-only positions
- why some staff do not engage with student evaluations
- how perceptions of student evaluations determine and influence teachers' beliefs and practices in their broader roles
- 'closing the evaluation loop' with students, including embedding 'closing the loop' notions explicitly in institutional evaluation/appraisal policy, practice and process
- perceptions of evaluation/appraisal held by teachers in tertiary education organisations other than universities and polytechnics
- views of other important groups who participate in evaluation/appraisal, namely, students and administration, management, and human resource personnel.

Table of Contents

<i>Key Words</i>	<i>i</i>
<i>Executive Summary</i>	<i>iii</i>
<i>Table of Contents</i>	<i>xii</i>
<i>List of Tables</i>	<i>xvii</i>
<i>List of Figures</i>	<i>xx</i>
<i>Acknowledgements</i>	<i>xxi</i>
Chapter 1 : Overview	1
1.1 Introduction	1
1.2 The Project	2
1.3 Significance	3
1.4 Intentions of the Study	3
1.5 Outline of Report Contents	4
Chapter 2 : Research Framework	6
2.1 Literature Review	7
2.1.1 Introduction	7
2.1.2 What are tertiary teachers' views of student evaluations?.....	8
2.1.3 What factors affect tertiary teachers' perceptions of student evaluation?	9
2.1.3.1 Perceptions of shortcomings in students' ability to evaluate their learning experiences	9
2.1.3.2 Institutional uses of student evaluations	10
2.1.3.3 Influence of teaching and learning beliefs on academics' perceptions of student evaluations	11
2.1.3.4 Emotional responses to student evaluation feedback	11
2.1.3.5 Perceptions that there need to be multiple sources of evaluation	13
2.1.4 How do tertiary teachers engage with student evaluations?.....	14
2.1.4.1 Challenges of interpreting student feedback.....	14
2.1.4.2 Support for more effective engagement with evaluations	15
2.1.4.3 A sense of personal agency to foster engagement with student evaluations.....	17
2.1.5 Summary of literature review.....	18
2.2 Environmental Scan	19

2.2.1 The New Zealand tertiary sector operating environment	20
2.2.2 Institutional evaluation systems	22
2.2.2.1 The process undertaken	22
2.2.3 Results of the scan.....	24
2.3 Setting - The Three Institutions Involved in the Project.....	37
2.3.1.1 The University of Otago.....	37
2.3.1.1.1 Background.....	37
2.3.1.1.2 Teaching	44
2.3.1.1.3 Teaching development.....	45
2.3.1.1.4 Other features of significance for evaluation/appraisal	46
2.3.1.2 The University of Waikato	48
2.3.1.2.1 Background.....	48
2.3.1.2.2 Teaching	49
2.3.1.2.3 Teaching development.....	50
2.3.1.2.4 Other features of significance to evaluation/appraisal	51
2.3.1.3 The Otago Polytechnic	51
2.3.1.3.1 Background.....	51
2.3.1.3.2 Teaching	52
2.3.1.3.3 Teaching development.....	54
2.3.1.3.4 Other features of significance to evaluation/appraisal	55
2.4 Summary of Chapter 2 – A Conceptual Framework	56
<i>Chapter 3 : Research Design.....</i>	59
3.1 Introduction	59
3.2 Aim, Research Questions and Objectives of the Study	59
3.3 Research Plan.....	60
3.4 Macro Research Design - Structure and Procedures	63
3.4.1 Participants and researchers	64
3.4.1.1 The tertiary teachers	64
3.4.1.2 The researchers.....	65
3.4.2 Data sources	66
3.4.2.1 Questionnaire	67
3.4.2.2 Interviews	69
3.4.3 Analysis techniques.....	71
3.4.3.1 Questionnaire	71
3.4.3.1.1 Quantitative data.....	71

3.4.3.1.2 Qualitative data.....	73
3.4.3.2 Interviews	74
3.5 Quality Assurance.....	75
3.6 Summary of Chapter 3.....	75
<i>Chapter 4 : Findings</i>.....	76
4.1 Introduction	76
4.2 The Demographics.....	76
4.3 The Questions.....	80
4.3.1 Question 1: What perceptions do tertiary teachers hold about student evaluations?.....	81
4.3.1.1 Teachers’ perceptions of student evaluations – questionnaire responses	81
4.3.1.2 Teachers’ perceptions of student evaluations – probing beliefs and emotions through the interviews	92
4.3.1.2.1 University of Otago	93
4.3.1.2.2 The University of Waikato	96
4.3.1.2.3 Otago Polytechnic.....	101
4.3.2 Conclusions	103
4.3.3 Question 2: What factors (causes, influences) affect these views?	104
4.3.3.1 Institutional requirements	105
4.3.3.2 Other institutional influences	107
4.3.3.3 Internal/external reporting requirements	113
4.3.3.4 Institutional culture and expectations	115
4.3.3.5 Evaluation/appraisal instruments and systems	116
4.3.3.5.1 Evaluation processes and instruments and institutional uses of evaluation – comments from the questionnaire.....	117
4.3.3.5.2 Evaluation processes and instruments and institutional uses of evaluation – comments from the interviews.....	122
4.3.4 Conclusions	129
4.3.5 Question 3: How do tertiary teachers engage with evaluation results and students' feedback?	129
4.3.5.1 Teacher engagement with evaluation – evidence from the questionnaire	130
4.3.5.2 Teacher engagement in evaluation – evidence from the interviews	151
4.3.5.2.1 University of Otago	151
4.3.5.2.2 The University of Waikato	153
4.3.5.2.3 Otago Polytechnic.....	156
4.3.6 Conclusions	157
4.4 Summary of Chapter 4.....	158

Chapter 5 : Discussion	159
5.1 Introduction	159
5.2 The Assertions	167
5.2.1 Assertion 1	168
5.2.2 Assertion 2	174
5.2.3 Assertion 3	180
5.2.4 Assertion 4	185
5.2.5 Assertion 5	190
5.3 Summary of Chapter 5	198
Chapter 6 : Conclusion	199
6.1 Summary of the Study	199
6.2 Recommendations	203
6.3 Further Investigation	211
6.4 Conclusion	213
Appendix 1 : Questionnaire	217
Appendix 2 : Research Notes Excerpt – Process of Identifying Prospective Questionnaire Respondents	224
Appendix 3 : Core Interview Questions	226
Appendix 4 : Project Timeline	228
Appendix 5 : Questionnaire Data	229
Appendix 6 : Comment Data Analysis Q10	242
Appendix 7 : Comment Data Analysis Q12	245
Appendix 8 : Comment Data Analysis Q14	249
Appendix 9 : Comment Data Analysis Q16	253
Appendix 10 : Comment Data Analysis Q18	257
Appendix 11 : Comment Data Analysis Q20	260
Appendix 12 : Comment Data Analysis Q21	264

<i>Appendix 13 : Comment Data Analysis Q22.....</i>	<i>268</i>
<i>Appendix 14 : Feeding Back to Students – An Example</i>	<i>272</i>
<i>References</i>	<i>275</i>

List of Tables

<i>Table 2:1 Institutional Websites Included in the Environmental Scan.....</i>	<i>22</i>
<i>Table 2:2 Summary of Key Features of Student Evaluation Systems in New Zealand Institutions.....</i>	<i>26</i>
<i>Table 2:3 Details of Environmental Scan, by Institution.....</i>	<i>28</i>
<i>Table 2:4 Vision, Mission and Values of Each Institution Involved in the Project...38</i>	
<i>Table 2:5 The Institutional Context Around Evaluation/Appraisal of the Three Participating Institutions.....</i>	<i>40</i>
<i>Table 3:1 Number of Tertiary Teachers Invited to Participate in the Study from across the Three Institutions.....</i>	<i>64</i>
<i>Table 3:2 Overall Distribution Report – Total Responses to Questionnaire</i>	<i>69</i>
<i>Table 3:3 Volunteers for Interview from each Participating Institution</i>	<i>70</i>
<i>Table 3:4: Summary of Response Rates – Comment Questions</i>	<i>73</i>
<i>Table 4:1: Overall Distribution Report – Total Responses to Questionnaire.....</i>	<i>76</i>
<i>Table 4:2: Summary of Participants’ Gender/Sex (Q23)</i>	<i>77</i>
<i>Table 4:3: Summary of Participants’ Ethnicity Data (Q24)</i>	<i>77</i>
<i>Table 4:4: Summary of Participants’ Age Data (Q25).....</i>	<i>78</i>
<i>Table 4:5: Summary of Participants’ Tertiary Teaching Experience (Q26).....</i>	<i>78</i>
<i>Table 4:6: Summary of the Nature of Employment of Participants (Q27).....</i>	<i>78</i>
<i>Table 4:7: Summary of Positions Held by Participants in their Institutions (Q28)..</i>	<i>78</i>
<i>Table 4:8: Percentage of Participants who had Applied/Not Applied for Promotion During the Previous Five Years (Q29).....</i>	<i>79</i>
<i>Table 4:9: Percentage of Participants who had been/had not been Promoted During the Previous Five Years (Q30).....</i>	<i>79</i>

<i>Table 4:10: Summary of Division/Faculty/School in which the Participants were Predominantly Based (Q31).....</i>	<i>79</i>
<i>Table 4:11: Summary of Location where the Participants were Predominantly Based (Q32)</i>	<i>80</i>
<i>Table 4:12 Frequency of Responses to Q17 (Do you personally think it is worthwhile to gather student appraisal/evaluation data about teaching and courses/papers?)</i>	<i>81</i>
<i>Table 4:13 Overall Responses to Q2 (Please identify why you use student evaluations/appraisals).....</i>	<i>86</i>
<i>Table 4:14 Overall Responses to Q4 (When you receive the results from your student evaluations/appraisals, do you...)</i>	<i>90</i>
<i>Table 4:15: Q4 Median Test Results</i>	<i>91</i>
<i>Table 4:16 Percentage of ‘Yes’ Responses to Q2a (because it is required by my department/school/institution)</i>	<i>105</i>
<i>Table 4:17 Combined Total Ratings for Q11 (To what extent does your institution’s use of student evaluation/appraisal data influence your teaching decisions?).....</i>	<i>107</i>
<i>Table 4:18 Significant Differences Between Tertiary Teaching Experience (Q26) and Q11 Responses</i>	<i>108</i>
<i>Table 4:19 Percentage of ‘Yes’ Responses to Q2h (to report on quality matters to relevant internal and external bodies)</i>	<i>113</i>
<i>Table 4:20 Combined Total Ratings for Q19</i>	<i>121</i>
<i>Table 4:21 Responses to Q17 (Do you consider it personally worthwhile to gather evaluation/appraisal data about teaching and courses/papers?).....</i>	<i>130</i>
<i>Table 4:22 Summary of ‘Yes’ Responses to Q2 with Proportion of Responses (Rankings 1–8) Made by Teachers at Each Institution.....</i>	<i>131</i>
<i>Table 4:23 Summary of Overall Ranked Responses to Q4a–4g.....</i>	<i>136</i>
<i>Table 4:24 Frequencies of ‘Yes’ Responses to Q6 and Q7</i>	<i>139</i>
<i>Table 5:1 Assertions</i>	<i>168</i>

Table 5:2 Assumptions and Possible Criticisms about Students as Sources of Data, Student Evaluation Systems, Processes and Instruments.....179

List of Figures

<i>Figure 2:1 Influences on tertiary teacher engagement with student evaluations/ appraisals</i>	57
<i>Figure 3:1 Research notes about the questionnaire pilot – an example</i>	67
<i>Figure 4:1 Perceptions of institutional influence on teaching decisions across institutions.</i>	108
<i>Figure 5:1 Factors influencing engagement with evaluation</i>	160

Acknowledgements

As with any research project, the work does not occur in isolation. We would like to acknowledge a range of people who have supported us as we travelled this journey of discovery. We have not listed people in any particular order, and if we have missed anyone out we are truly sorry, but we wish to thank you all very much for your assistance, guidance, advice and support.

- Ako Aotearoa and the National Project Fund for funding this piece of work, and more specifically Dr Kirsty Weir, Dr Peter Coolbear and Nyk Huntington for their support and enthusiasm about the project.
- Members of the senior leadership teams at the three institutions for supporting us and the questionnaire.
- Carla Thompson at the University of Otago for her wonderful work on the annotated literature review, which formed the basis of the project.
- Pip Bruce Ferguson at The University of Waikato for interviewing questionnaire respondents and providing advice about the presentation of the final report.
- Leen Al-Hadban at The University of Waikato and Angela French at the University of Otago for help with transcribing the interviews.
- Professor John Harroway at the University of Otago for advice about the form and structure of the questionnaire and statistical analysis of the questionnaire data.

And last but not least in our thoughts, the respondents to the questionnaire and our interviewees. This project could not have been completed without your involvement.

Thank you all,

The research team.

Chapter 1 : Overview

1.1 Introduction

This study was prompted by an interest in evaluation, and in particular, the centralised student evaluation system that virtually all tertiary institutions have in place. The research team are from three different tertiary institutions with roles that focus on the administration of student evaluations/appraisals or on broader aspects of academic staff development. These roles involve working with staff, including teachers, heads of schools and departments and other management personnel, and often include interactions concerning evaluation and the use and interpretation of results of student evaluation questionnaires. Teachers can find 'poor' evaluations confronting and be upset emotionally by the reports they receive. Some fear that their jobs may be in jeopardy because of these 'poor' results.

From an academic staff development point of view, evaluation is a core aspect of teaching practice and engaging in and with evaluation processes is an essential component of reflective practice. Evaluation should not be a punitive measure of effectiveness. A centralised evaluation system in an institution can provide a service to teachers and departments as they engage in continual cycles of planning and refining their teaching and learning beliefs and strategies. Evaluations can help academics to try new ideas, gauge the success of their teaching and courses and review the experiences of students.

However, the research team acknowledges that institutions are interested in, and indeed, have a right and responsibility to know about, the activities of their staff, and are therefore keen to know that staff are performing effectively. Institutions are also responsible for providing support for staff who experience difficulties with their teaching and for ensuring that professional development is available for their teachers to enhance already good teaching practices.

The concern of this study was the institutional use of student evaluation questionnaires as a means of gauging the effectiveness of teachers and teaching in ways that are, anecdotally, seen as punitive. The focus of the research was to investigate whether staff perceive the use of the same instrument for quality as well as developmental purposes as intrinsically contradictory and unhelpful.

The topic of student evaluations has received a great deal of attention in the literature and while much has been written and researched, relatively little is documented about the New Zealand tertiary education scene. From working in the areas of student evaluation and/or academic staff development it was evident to the members of the research team that professional development uses of summative student evaluation tends to be limited, and students frequently do not get the opportunity to see the outcomes of the feedback they provide.

By studying the topic in this project, the intentions were to gain some insight into how evaluations are having an impact on academic staff in New Zealand institutions with a view to setting the scene for change. It was hoped that the study would provide some indication of the nature and cause of the emotional and personal responses that are encountered regularly by teachers and provide some understanding of the evaluation situation that is based on evidence, not on anecdotes and perhaps even myths. The intention was to undertake a New Zealand-based study to increase the relevancy of the research. The long-term goal was to contribute to the evolution of student evaluation systems that are seen and accepted as useful and informative by all involved in the educative process – teaching staff, management and students – and that can improve students' learning experiences.

1.2 The Project

Ako's vision includes "the best possible education outcomes for all learners" (Ako Aotearoa, n.d.). Research in higher education argues for the importance of aligning teaching practices as closely as possible with the needs and approaches of learners. The 'standard' student evaluation of teaching/course questionnaires administered centrally by institutions can provide a broad picture of students' experiences of learning and therefore inform teachers about the effectiveness and quality of both teaching and learning. However, student evaluations are also a widely-used institutional quality mechanism, often used in staff promotion processes. It is acknowledged that the accountability and developmental purposes of student evaluations are both pertinent in contemporary tertiary education. However, student evaluations in tertiary institutions are primarily used for gauging and demonstrating effectiveness of teaching, suggesting that their potential for improving teaching and students' learning experiences is undervalued and underutilised.

As explained in the previous section, anecdotal evidence suggests that the tensions that can be perceived between the two purposes may lead to suspicion about negative attitudes to student evaluations as well as unwillingness to maximise their professional development potential. With a view to confirming or refuting this anecdotal evidence, this study gathered empirical data about perceptions of evaluation held by tertiary teachers in New Zealand. The study also explored implications of those perceptions for teachers' practices and their ways of working in institutions that use evaluations to gauge the effectiveness of their teaching.

This document is the report of this study, which was funded through Ako Aotearoa's National Project Fund 2009.

1.3 Significance

This study explored the extent to which teacher thinking and practices are currently informed by learner feedback via evaluations. The exploration of the influence of evaluations on teacher professional development and student learning means that the study was implemented with a view to contributing directly to Ako Aotearoa's National Project Fund 2009 priority area of Strategic Professional Development Initiatives. Recommendations are made to enable teachers and institutions to optimise feedback gained through student evaluations and align activities, whether related to development of teaching or accountability, more effectively with learners' needs.

1.4 Intentions of the Study

This study aimed to consolidate, evaluate and add to past studies of student evaluations of teaching by investigating, from a New Zealand perspective, the extent to which teacher thinking and practices are currently informed by learner feedback.

The research question that provided focus for the study was:

How do the current evaluation processes and practices influence teachers' thinking and behaviours in relation to student learning at all stages of the teaching and learning cycle?

Further questions to uncover some of the key issues contributing to answering the major question included:

- *What are the perceptions that tertiary teachers hold about student evaluation/appraisal?*
- *What factors (causes, influences) affect these views?*
- *How do tertiary teachers engage with student evaluations/appraisals?*

With these research questions in mind, the project was planned around a series of objectives. The objectives provided the framework for the practical implementation of research activity including the data gathering, analysis and synthesis. The objectives were to:

- explore tertiary teachers' perceptions about student evaluations and identify how those perceptions have an impact on their thinking and practices
- explore how tertiary teachers make use of information obtained from student evaluations at all stages of the teaching and learning cycle
- compare teacher thinking and behaviours around evaluations across three participating institutions
- make recommendations about evaluation processes so that evaluations can be used optimally for teaching development and enhancing student learning.

In this report, the terms 'evaluations' and 'student evaluations' refer to the centralised student evaluation of teaching/course systems processes and instruments used by tertiary institutions to gather student feedback on teaching and/or courses. Because 'appraisal' is the term used at one of the institutions participating in the study (The University of Waikato), throughout this report 'evaluation/appraisal', 'evaluation' and 'student evaluation' are used interchangeably to accommodate the different usage in the institutions.

1.5 Outline of Report Contents

This report, which is a full description of the study, is made up of a number of chapters. Chapter 1 is an overview of the study. Chapter 2 presents the research framework comprising a review of the literature, a scan of the evaluation policies of New Zealand tertiary institutions and background material describing the contexts of the three institutions that participated in the study. Chapter 3 provides an overview of the research design, including descriptions of the research approach taken for the study, the study participants, data sources, analysis techniques and quality assurance measures. Following this, Chapter 4 presents the

findings around the three research questions, while Chapter 5 builds on the discussion of the findings around a set of five assertions. Finally, Chapter 6 summarises the report and presents a set of recommendations for practical action and future research.

Chapter 2 : Research Framework

A number of issues related to evaluation have been raised in the literature. These issues, borne out of experiences, views and beliefs held by various individuals and groups about students, teachers and teaching, and about evaluation questionnaires and systems, link with the many evaluation myths and anecdotes that exist in a tertiary education environment.

Topics related to students and evaluations include concerns about students not providing worthwhile information via evaluation questionnaires, for a number of reasons:

- that students are ‘over-evaluated’, suffer from ‘survey fatigue’ and therefore do not provide quality information
- that students have an underlying fear of providing honest responses because teachers hold power over them, particularly where assessment is concerned
- that students are basically lazy, so not only are they not willing to think about their learning experiences in a way that could be helpful to their teachers, if the subject is difficult and they have been forced to work hard, they will naturally respond to evaluation questionnaires negatively
- that for most students, evaluation is simply a popularity contest, meaning students are not concerned about good teaching and give the teachers who are the most ‘fun’ or the ‘easiest on them’ the highest evaluation scores.

These views, related to teachers’ views about the capability of students to provide quality information, are bound up with deeper suspicions about the quality of the evaluation systems and instruments. Not only are the data from students seen as being unreliable, but the questionnaires themselves are perceived as inadequate and unsatisfactory measures of teaching. Measuring the quality of teaching according to a set of numbers and making decisions about a teacher’s future on the basis of the result of one question within the questionnaire is seen as inappropriate. These strong doubts about the inadequacy of evaluation systems result in teachers feeling constrained about their teaching, either because they are receiving good evaluations and do not want to change or develop their practices when clearly students are positive about their courses, or because they are afraid that introducing innovative changes might backfire and cause a negative response from students.

Critics of tertiary teachers claim that teachers view teaching simply as ‘content’ and do not engage in systematic review of their practice, and so centralised evaluation systems are a way to engage teachers in some way. Critics of centralised student evaluation systems on the other hand, hold issue with such systems, blaming them for reducing the notion of evaluation as part of reflection on practice, to a minimalist and very narrow one that promotes an emaciated image of evaluation.

Overriding all of the above, literature has dealt with the impact that institutions have on teaching; to value or not to value teaching over research and other academic activities. Expectations of institutions about teaching are reflected in the nature of rewards and incentives. For many teachers, this view results in their expending effort on either developing themselves as teachers or concentrating more on research.

Through this project, an intention was, therefore, to examine the literature in a systematic way, to identify the issues more succinctly, and to map out how researchers have come to grips with the many factors influencing and determining the nature and form of student evaluation in tertiary institutions. This review appears in section 2.1. In addition, a scan of tertiary institutions in New Zealand was undertaken to gain insight into how centralised student evaluations are managed in practice. In particular, a close examination was made of the three institutions that participated in the study. Thus, the theoretical and practical background was laid out. This environmental scan of New Zealand’s tertiary institutions appears in section 2.2, and the more in-depth detail about the institutional settings of the participating institutions appears in section 2.3.

2.1 Literature Review

2.1.1 Introduction

The subject of student evaluations of teaching is one of the most widely investigated topics in higher education research (Benton & Cashin, 2012). The concept of student evaluations of teaching and the debates about the merits of these processes can be traced back as early as the 1920s (D’Apollonia & Abrami, 1997; McKeachie, 1990; Menges & Mathis, 1988; Millman, 1981; Smock & Crooks, 1973). While earlier scholarship focussed on matters of validity and reliability, both of the instrument and of student evaluators (Beran, Violati & Kline, 2007;

McKeachie, 1990), more recently, there is a growing body of scholarship that examines the impact of evaluation data on teaching and learning processes for all involved.

Contextual factors have contributed to current scrutiny of the relationship between student evaluations and performance improvement. A key trend that has been widely identified in the literature is the proliferation of quality initiatives that are impinging on university teaching and which elevate student feedback to a new level of importance in measuring and rewarding tertiary teaching performance (Arthur, 2009; Moore & Kuol, 2005). In this climate it becomes increasingly important to ensure that the teaching and learning potential of student feedback does not become subsumed in the quality agenda. Instead, processes need to be explored that can enable the developmental possibilities of student evaluations to flourish alongside the quality initiatives. The backdrop of intensive quality initiatives must also be remembered when exploring academics' conceptions of student evaluations and their response to feedback. It may be that the quality framework is in itself a stimulus for negative perceptions, bearing in mind that academia is an environment "in which independent action and academic freedom is fiercely protected" (Moore & Kuol, 2005, p. 58).

This review critically outlines the main research trends and findings in the literature about teacher perceptions of student evaluations, in order to locate the current study and argue for its particular emphasis. The present study investigated the impact of teachers' perceptions of the student evaluations system on teacher thinking and behaviours at all stages of the teaching and learning cycle. The literature is grouped in relation to the three key research questions:

What are tertiary teachers' views of student evaluations?

What factors affect tertiary teachers' perceptions of student evaluations?

How do tertiary teachers engage with student evaluations?

2.1.2 What are tertiary teachers' views of student evaluations?

There is a widely held perception that academics are hostile to student evaluations. However, some research studies have countered this widely reported view. Schmelkin, Spencer and Gellman (1997) concluded that teachers' attitudes to the overall usefulness of student evaluations were positive, while Nasser and Fresko (2002) found that the teachers in their study were "mildly positive" about student evaluations. Braskamp and Ory (1994) also refute

many of the common concerns about the problems associated with student evaluations and the notion of a more positive faculty outlook is supported by the studies of Penny and Coe (2004) and Beran and Rokosh (2009).

2.1.3 What factors affect tertiary teachers' perceptions of student evaluation?

The considerable research on the validity and the reliability of the student evaluation process and instruments (Beran et al, 2007) was not the focus of the current study, but some of the main themes need to be documented as teacher perceptions of the student evaluation instrument and process are likely to be influenced by their beliefs about its validity and reliability. As noted by Costin, Greenough and Menges (1971), the “uses to which the student rating results are put depend heavily on faculty confidence in their meanings” (p. 529). The perceptions uncovered here are considerably varied and range from strong support (McKeachie, 1990) to a belief that student ratings are reasonably valid (Beran & Rokosh, 2009).

2.1.3.1 Perceptions of shortcomings in students' ability to evaluate their learning experiences

There is considerable evidence that student feedback questionnaires validly and reliably indicate the quality of teaching and learning (*e.g.*, Kember, Leung & Kwan, 2002; Marsh, 1987; Theall & Franklin, 2001). Other studies, however, summarise a raft of teachers' perceptions about the potential unreliability of student evaluation data, although many of these are based on reported opinion rather than empirical evidence. Aleamoni (1981) synthesised some teachers' common concerns that have been reported in the literature. These include the view that students lack the maturity to evaluate the quality of teaching and that their limited subject knowledge impedes their capacity to pass judgement. A further reported concern is that students are not in a position to assess the effectiveness of the teaching until a passage of time has elapsed. Other misgivings relate to the idea that irrelevant variables influence students' perceptions of the merits of a course and the teaching. These include factors such as the difficulty of a course, the grading habits of the teacher and the more general popularity of the teacher.

2.1.3.2 Institutional uses of student evaluations

One of the initial premises of the current study was that there may be a tension between institutional requirements and uses of student evaluation processes and their potential to inform teaching improvement. This tension is reported on, or speculated about, in some of the literature. Costin ,(1971) suggested that “faculty confidence” in evaluation/appraisal is linked to their use and that undiscerning use of evaluation instruments by institutions may undermine academics’ faith in their meaningfulness. McKeachie (1997), who argued persuasively that student ratings need to be respected, suggested that “the problem lies neither in the ratings nor in the correctness but rather in the lack of sophistication of personnel committees who use the ratings” (p. 1218). Penny and Coe (2004) say that academic suspicion about student evaluations is linked to the perception that they are primarily used as a quality monitoring instrument. This perception is heightened by the fact that student evaluations are generally not part of an integrated system of consultation and support, which would reinforce their teaching and learning potential.

Some writers speculate that the uses of student evaluation for formative development may be constrained by possible mistrust “because it is intended to serve both formative and summative purposes” (Beran & Rokosh, 2009, p. 182). A similar finding is reported by Edström (2008) who argued that the use of the same instrument for audit and developmental purposes undermines the potential usefulness of student evaluations as a tool to improve teaching. The study of Nasser and Fresko (2002) in an institution where evaluation had originally been used only for formative purposes, revealed that few faculty members were in favour of results being sent to administrators. The tension between institutional and developmental purposes of student evaluations is the hypothesis underlying the study of Arthur (2009), who argues that centrally administered audit-focussed evaluation systems appear to destabilise academics’ sense of professional autonomy and willingness to use them to enhance their teaching. Moore and Kuol (2005) note that mistrust of organisational use of evaluations is a factor contributing to the negative perceptions that some academics have of appraisal/evaluation and that it is important to acknowledge the shift in the “balance of power within [tertiary] institutions” (p. 59) to students if student evaluations influence academics’ rewards and career progression. Writing in the context of the Irish tertiary sector, where compulsory student evaluations had not previously been the norm, Moore and Kuol suggest

that this shift of authority to students “may indeed be the reason why so many arguments against their introduction have reached both public and scholarly arenas” (p. 59).

2.1.3.3 Influence of teaching and learning beliefs on academics’ perceptions of student evaluations

Undoubtedly, there is a range of factors that help to explain the gap between teachers’ acquiescence to the notion of student evaluations and their limited use of the feedback to inform teaching improvements. Teachers’ receptivity to, and use of, student evaluation feedback may be influenced by their teaching and learning beliefs, including conceptions about the nature of knowledge and their views of power relationships in the teaching and learning context. This connection has not been explored extensively in the evaluations literature, but a study by Hendry, Lyon and Henderson-Smart (2007) specifically investigated this possible relationship. Participants in this study were teachers on a problem-based medical programme at the University of Sydney. All lecturers and presenters in this programme over a two-year period were asked to anonymously complete the *Approaches to Teaching Inventory* (Prosser & Trigwell, 1999) and an *Approach to Feedback Inventory* (Hendry *et al.*, 2007). Results suggested that the types of changes teachers made to their courses aligned with their conceptual views. The study indicated that teachers with an approach that was student-focussed and who saw learning as involving strong conceptual change were more responsive to feedback and more positive about strategies for improving their teaching. By contrast, those teachers who were strong on an information-transmission approach found interpreting feedback difficult. The authors argue that education of teachers to improve their use of student feedback needs to be integrated into a wider “systematic programme of teacher education that focuses on teachers’ conceptions, approaches and supports observation of exemplary practice” (Hendry *et al.*, 2007, p. 152).

2.1.3.4 Emotional responses to student evaluation feedback

While academics’ views about student evaluation have been fairly widely reported, although not always supported by evidence, there has been only minor consideration of the emotion generated when reading student feedback. Furthermore, little research attention has been given to the impact that this emotion may have on academics’ subsequent reading and use of feedback to inform changes in their teaching. It has been noted in this review that little

attention has been devoted to the terrain between a reported, if often reluctant or half-hearted, acquiescence to student evaluations and the relatively limited actual use of student feedback for teaching improvement (Beran & Rokosh, 2009; Nasser & Fresko, 2002; Schmelkin *et al.*, 1997). One component of this issue could be the subjective emotional responses that academics undergo when reading student feedback, which may, in turn, colour or limit its usefulness for professional development and performance improvement. This gap in the literature has been noted in more recent literature on student evaluations, which have started to explore the feelings generated on reading student feedback. Moore and Kuol (2005) suggest that this is a significant lacuna in the literature as “the individual reaction to performance feedback has a more direct bearing on any subsequent efforts to improve, sustain, enhance or develop performance in the future” (p. 61). Moore and Kuol (2005) concluded that their preliminary research findings confirmed a link between feedback reaction and subsequent teaching performance enhancement. Some of their interesting results included that while 44 per cent of academics’ largely positive responses correlated with students’ largely positive comments, 27 per cent of the surveyed academics focussed on negative information from largely positive comments. There was also a surprising gap between academics’ expectations of their students’ responses and what the students actually said. Fifteen academics of the 21 in the study reported being surprised by students’ feedback, with slightly more being surprised by negative than positive comments (six were surprised by positive feedback, while nine were surprised by negative feedback).

Moore and Kuol’s (2005) study was a significant one for the current research because of the shared focus on lecturers’ perceptions, in their case primarily on ‘feedback reaction’. The present study was interested in the emotional dimension of lecturers’ responses to performance feedback, but investigated this aspect as one of the many factors potentially shaping lecturers’ perceptions and use of student evaluation data. There are also some questions about the potential transferability of Moore and Kuol’s findings. As the authors themselves noted, the study was a small scale exploratory study that did not include ‘face-to-face’ encounters. Furthermore, the study was conducted in a very different context from the study presented in this report, one in which student evaluations of teaching had not been the norm.

Another small scale study into lecturers' responses to student feedback was conducted by Arthur (2009). Arthur comments that while there is evidence that emotions play a part in student learning, little attention has been given to the emotions of lecturers in the process of developing skills, particularly to the significance of "their affective response to student evaluations" (Arthur, 2009, p. 444). This study, which was based on semi-structured interviews with eight university lecturers, took up the challenge issued by Moore and Kuol (2005) for more face-to-face encounters to explore the affective aspect of lecturers' responses to student feedback. A key conceptual framework for Arthur (2009) was the perception that the student evaluation process sits at the interface of performativity and professionalism and that this coexistence has important implications for the way academics respond to, and act on, student evaluations of teaching. Arthur (2009) found that "all the interviewees expressed emotional responses to student feedback" (p. 449). They found that this was particularly true for academics receiving negative student feedback in the early stages of their career, but even more seasoned academics demonstrated sharp emotional responses to critical feedback. As Arthur observed, these emotional reactions are related to the strong link between the personal and the professional in a teacher's work.

The current study, across three institutions, included an interview component in which the researchers aimed to elicit "rich discussion about teachers' emotional and pragmatic responses to student feedback" (Moore & Kuol, 2005, p. 70). Although the focus of the study was not on emotional responses to student feedback, the interview component provided the opportunity for emotions to be reported and expressed and thus extended the interview process conducted on a small scale by Arthur (2009).

2.1.3.5 Perceptions that there need to be multiple sources of evaluation

While there seems to be moderate acceptance of the use of student evaluations in tertiary institutions (Beran & Rokosh, 2009), this is often passive and does not readily translate into re-evaluation of, or substantial changes to, teaching. As Hendry *et al.*, (2007) observed, the original meaning of feedback, which derives from the sciences as an "automatic modification of a process or a system by its effects" cannot be assumed in the context of student evaluations of teaching (p. 143). One possible explanation for the limited use of student evaluations to inform teaching may be the widely voiced perception that student evaluations of teaching should be seen and interpreted within the context of other sources of evaluation

such as peer and self-review. For example, Braskamp and Ory (1994) found that most faculties see student evaluations as one important indicator of teaching quality, while Penny and Coe's (2004) study showed that faculties want student evaluation findings to be allied with other forms of feedback such as that of peer and self-evaluation, a result endorsed by Ballantyne, Borthwick and Packer (2000) and Burden (2008). Interest in other complementary sources of evaluation may also be associated with one recurrent truism in academia: the perception that students are not in a position to make judgements about the quality of some aspects of teaching, such as specialist knowledge (Aleamoni, 1981).

2.1.4 How do tertiary teachers engage with student evaluations?

While the current literature indicates relatively positive attitudes towards student evaluations, which appears to contradict the more commonly reported academic hostility towards the instruments, Beran and Rokosh (2009) and Burden (2008) remind us that the acceptance of the student evaluation scheme does not correlate with perceptions of its usefulness to enhance teaching, or with actual usage of the instrument for teaching changes. Beran and Rokosh (2009) speculate that "since instructors find ratings to be of little practical value, their seemingly positive attitudes regarding student ratings actually reflects a neutral view point or a passive acceptance of the ratings in general" (2009, p. 183). Similarly, Smith (2008) noted that there is "little published evidence that they [evaluations] are systematically used for developing and improving their teaching" (p. 518). Gauging the terrain between intellectual acceptance of student evaluations of teaching and tertiary teachers' actual use of appraisals/evaluations to inform teaching development was one of the questions that the current study probed. As Ory and Ryan (2001) contended, unless this failure to use evaluations to inform teaching and develop processes that help to remedy this can be understood, the student evaluations process will be little more than a ritual that both teachers and students participate in because it is compulsory.

2.1.4.1 Challenges of interpreting student feedback

As Kember *et al.*,(2002) report, running student evaluation questionnaires does not automatically lead to improvements in teaching. While a number of studies indicate the usefulness of student evaluations for teacher development, this promise is undermined by the fact that the evaluations process is often an isolated exercise and that institutions generally

provide very little guidance to staff in relation to the interpretation of student evaluations (Arthur, 2009). Penny and Coe (2004) cite the findings of Cohen's (1981) review, which argued that students' ratings on their own were not enough to facilitate teaching improvements and that they needed to be supplemented by a consultation process. Penny and Coe (2004) outlined key elements that they believe should be part of a consultation process. They proposed that consultation with teachers around student evaluations and possible teaching improvements work best when other sources of evaluation are incorporated into the discussion. Specifically, they highlighted the usefulness of incorporating self-ratings and peer feedback on teaching into the consultations.

The importance of reading student evaluations in the context of multiple sources of evidence is extended by Smith (2008) who proposes a "Four Quadrant" approach to evaluation, with the quadrants including self-reflection, peer review, student learning and student experience. Smith (2008) contended that this "quadrangulation" corresponds to the use of triangulation in research, which means "to gather data on a phenomenon from a variety of angles or perspectives or sources to give a better more accurate picture of the thing you are studying" (p. 528). Likewise, drawing on multiple sources in the context of teaching evaluation can make feedback both "more believable and more empowering for academics" (Smith, 2008, p. 528). Smith (2008) also suggested that an interpretative guidance system could help to combat the weak correlation between student evaluations and subsequent improvements to teaching. He observed that "there is little published evidence that they are systematically used by staff for developing and improving their teaching" (p. 518). His study outlines the development of interpretative guidelines as one part of a "five-phase programme linking evaluation and staff development" (p. 520).

2.1.4.2 Support for more effective engagement with evaluations

In their discussion of research conducted on lecturers' attitudes to student evaluations at a large Canadian university, Beran and Rokosh (2009) found that "over half agreed or strongly agreed that the feedback provided by the student ratings instrument is useful", yet, "only a few instructors reported having made any substantial modification to individual aspects of instruction as a result of student feedback" (p. 182). In trying to account for this disparity the authors concurred with Centra (1993) that the most significant impediment to teaching improvement is that teachers do not know how to make the appropriate changes to their

practice. In order to convert student evaluations into an integral and normal element of professional development, appropriate institutional systems need to be put in place.

Ballantyne *et al.*,(2000) report on a project developed at Queensland University of Technology in which selected academic staff and students were surveyed about their priorities for academic development. Groups consisting of both staff and students then collaborated to design developmental booklets that responded to these views. One of the very attractive aspects of this project was the collaboration of staff and students, as evaluation processes are generally characterised by a gulf between reports of student learning experiences and teachers' reception of this information. It is possible that this chasm, whether perceived or real, may help to explain the levels of animosity that both sides may feel towards the evaluation process.

Another small scale example, by Fisher and Miller (2008), illustrates the effects of changing thinking about student evaluation from involving “more traditional approaches that focussed on one collection of retrospective feedback in an end-of-course numerically based teaching evaluation” to “an ongoing and iterative process” (p. 200). Such a change can result in using data collection instruments differently and gaining more from any student evaluation activity than simply a series of qualitative results. Their study involved tutors and students as contributors to the evaluation process and incorporated student reflections on their own learning, as well as on the features of the learning environment that were having an impact on that learning. In a similar, yet more informal vein, Bovill (2011) discussed the impact of thinking about evaluation *for* learning rather than *of* learning in her study, and argued for greater and more meaningful involvement of students in evaluation.

As already mentioned, Smith (2008) proposed a detailed model for integrating student evaluations and academic development. Smith argued that for institutional purposes, managers generally focus on aggregate data, but improvements that really enhance the students' learning experience need to happen at the individual teacher level. This is supported by the comments of Penny and Coe (2004) who argued that supporting consultation needs to be context and person specific. Smith's five phase model includes evaluations, interpretive guidance, compiling of reports that indicate common patterns in evaluations of teacher performance and subsequent invitation to appropriate professional development. Professional development includes introduction to the “4Q” evaluation framework within the context of

learning communities, use of an action learning approach and workshops, as well as individual consultancies. The model aims to move practitioners towards “increasingly independent engagement, in the scholarship of teaching” (Boyer, 1990). Smith’s model offers a comprehensive model for optimising learning from the student evaluations process and translating the findings into enhancement of practice. The inclusion of the concept of a community of learning also has the additional advantage of promoting cultural change around teaching through foregrounding teaching-related conversations. However, as Smith observed, the integrated and comprehensive nature of this model needs both institutional support and extensive resourcing, but it is probably the only kind of approach that can transform student evaluations from a ritualistic process to an informative source for teaching enhancement. Furthermore, in order to promote serious uptake of the developmental opportunities, it is likely that institutions need to provide incentives and rewards for evidence of teaching improvement (Nasser & Fresko, 2002).

2.1.4.3 A sense of personal agency to foster engagement with student evaluations

The models of Ballantyne *et al.*,(2006) and Smith (2008) both provide academic staff members with a degree of personal agency in the interpretation of evaluations results and associated professional development. Ownership is arguably a key component for any system that is designed to integrate student evaluation and professional development more effectively. As Arthur (2009) suggested, academics are “less likely to act on the findings of student feedback if it is collected and analysed centrally (for performativity purposes) because this divorces the findings from the context of teaching and learning” (p. 443). Arthur argued that a “performativity culture” reduces academics’ sense of control and influence, and lecturers are more likely to see student evaluations as an imposed ritual. According to the typology he developed, academics in this culture are more likely to respond to negative evaluations with a “blame the students” reaction. This is because, Arthur contends, a performativity culture removes academics from agency for their own teaching. By contrast, according to Arthur, in cultures that emphasise academic professionalism, academics are more likely to modify aspects of their practice (‘tame’) or re-evaluate it. Discussions about interpretation of, and appropriate support for, subsequent development must therefore be conducted within the broader framework of the institutional use of evaluations data. The

question of ownership or perceived ownership of the student evaluations process emerges from the literature review as another telling question that the current research study explored.

2.1.5 Summary of literature review

This chapter began with an outline of the impetus for the study, the goal to examine academics' perceptions of student evaluations, so as to investigate ways of maximising the potential of student feedback for enhancing teaching and student learning. The literature review mapped out the territory of concern and was organised to reflect the variety of elements influencing and determining teachers' evaluation thinking and practices. Factors such as perceptions about students' capacity to make judgements about teaching, the institutional use of evaluation data, emotional responses to student feedback and the limits of the instruments including the need for multiple sources of data, were included in the discussion. In addition, more specific notions related to teacher engagement with, and ownership of, student evaluation processes were highlighted, including the issue of bridging the gap between gathering data for institutional purposes as well as for professional development. In doing so, the review brought to the fore the many challenges related to student evaluation/appraisal that educationalists have identified and documented in the past.

In summary, the literature review has highlighted the following important aspects of teachers' evaluation views and practices:

(i) Perceptions of evaluation/appraisal

The literature suggests that there is a widely reported view that academics are hostile towards student evaluations. However, more recent research argues that, instead, academics are generally resigned to the notion of student evaluations as a fact of the contemporary tertiary environment. Even so, the research concurs that this notional acceptance does not translate into serious engagement with student evaluation feedback as a tool for professional development.

The current study aimed to examine these propositions about teacher perceptions of evaluations in a New Zealand context and investigate their impact on teacher thinking and behaviours.

(ii) Factors influencing those views and practices around evaluations

The literature suggested that a number of interrelated factors influence academics' perceptions of evaluations and their use of them. Previous studies indicate that factors such as institutional expectations and community norms can influence attitudes to evaluation/appraisal and consequent responses. These factors include teachers' perceptions of the limitations of student judgement, the quality of their institution's evaluation instruments, the institutional ownership and use of evaluations, and the individual teacher's teaching beliefs and emotions.

The current study sought to investigate the relative significance of these factors for New Zealand tertiary teachers in terms of their views of, and engagement with, student feedback in evaluations. Correspondingly, the research investigators aimed to recommend ways of addressing academics' issues in relation to these factors.

(iii) Teachers' engagement with evaluations/appraisals processes and systems

The literature suggested that there are contextual, philosophical, practical and personal factors that influence the way academics engage with, and utilise, feedback from student evaluations. This study sought to investigate the influence of these factors in a New Zealand context, so as to recommend processes that can enhance engagement and thereby improve the student learning experience.

These three aspects, drawn from across the literature review, contributed to the refinement of the research questions and the shape of the practical implementation of the study. The next sections (sections 2.2 and 2.3) will illustrate how tertiary institutions in New Zealand are attempting to translate broader evaluation notions, ideals and principles about student evaluation into practical and workable ways within their own communities through their policies and processes. The following environmental scan works alongside the literature review to provide a backdrop against which the current study was undertaken.

2.2 Environmental Scan

As explained at the start of this chapter, as well as surveying the literature, it was deemed important to investigate the ways in which student evaluations were viewed and presented in New Zealand tertiary institutions. This section presents an overview undertaken via a scan of

websites of a variety of public documents. First, the expectations of tertiary institutions in terms of accountability to government as stated by the New Zealand Tertiary Education Committee (TEC) are outlined. This is followed by a summary of the public documents about evaluation policies and practices from each of the tertiary institutions included in the scan.

2.2.1 The New Zealand tertiary sector operating environment

The New Zealand tertiary sector is made up of a number of types of providers, namely, universities, polytechnics, wānanga, private training establishments, industry training organisations and adult and community education providers. The *Tertiary Education Strategy 2010 – 2015* outlines the roles and expectations of each group of providers. The literature review indicated strongly that context – both the immediate teaching setting as well as the broader institutional setting – plays an important part in influencing and shaping teachers’ perceptions of evaluations. Government expectations will cause institutions to concentrate their efforts and resources in ways that will meet those demands, resulting in various institutional policies, processes and practices. The nature of an institution, will, in turn, ‘flavour’ the policies, processes and practices, people, history and culture. This means that although there will be commonalities across all tertiary institutions, and then across all similar institution types, each institution will have a unique character.

The setting for the current study, which aimed to examine teachers’ perceptions of student evaluations, was within two universities and one polytechnic. To provide some broad information about the demands and expectations of those two types of tertiary education providers, the expectations and roles of universities and polytechnics as described in the *Tertiary Education Strategy 2010 – 2015*

(<http://www.minedu.govt.nz/NZEducation/EducationPolicies/TertiaryEducation/PolicyAndStrategy/TertiaryEducationStrategy/PartThreeExpectations.aspx>) are now summarised.

The *Tertiary Education Strategy 2010 – 2015* states that universities have three core roles. Which are to:

1. Undertake research that adds to the store of knowledge.
2. Provide a wide range of research-led degree and post-graduate education that is of an international standard.

3. Act as sources of critical thinking and intellectual talent.

The emphasis is on research, knowledge creation, and the provision of degree and higher degree programmes. In addition, the Government expects universities to:

- enable a wide range of students to successfully complete degree and post-graduate qualifications
- undertake internationally-recognised original research
- create and share new knowledge that contributes to New Zealand's economic and social development and environmental management.

Access to education for a wider range of students, international recognition, knowledge creation and dissemination to enhance the country's economy, society and environment are the major thrusts.

Polytechnics, on the other hand, are given a different emphasis. Their three core roles are to:

1. Deliver vocational education that provides skills for employment.
2. Undertake applied research that supports vocational learning and technology transfer.
3. Assist progression to higher levels of learning or work through foundation education.

A focus on vocational education, applied research and support for learners to experience higher levels of learning are the drivers behind the work of polytechnics. Following this, the government expects polytechnics to:

- enable a wide range of students to complete industry-relevant certificate, diploma and applied degree qualifications
- enable local access to appropriate tertiary education
- support students with low literacy, language and numeracy skills to improve these skills and progress to higher levels of learning
- work with industry to ensure that vocational learning meets industry needs.

There is a concentration on certificate, diploma and applied degree programmes, local access and industry collaboration and cooperation, with an explicit provision of literacy and numeracy learning support.

This variation in demands placed upon universities and polytechnics forms the basis of the different expectations that individual institutions have of their staff, which in turn influences each institution's evaluation/appraisal system.

2.2.2 Institutional evaluation systems

The purpose of this section is to present an overview of the New Zealand scene in terms of the public statements that tertiary institutions have made about the role played by student evaluations/appraisals. This builds on the wider roles and expectations of tertiary institutions as described in the *Tertiary Education Strategy 2010 – 2015* and provides a further contextual basis to how the institutions that participated in the study are placed within the broader New Zealand tertiary sector where evaluation is concerned.

2.2.2.1 The process undertaken

In April 2010, an online search was carried out in New Zealand's eight universities' and 20 polytechnics' official websites. Table 2:1 lists the institutions included in the environmental scan.

Table 2:1 Institutional Websites Included in the Environmental Scan

Institution type	Institution name
Polytechnics	Aoraki Polytechnic
	Bay of Plenty Polytechnic
	Christchurch Polytechnic Institute of Technology
	Eastern Institute of Technology
	Manukau Institute of Technology
	Nelson Marlborough Institute of Technology

	NorthTec
	Otago Polytechnic
	Southern Institute of Technology
	Tai Poutini Polytechnic
	Tairāwhiti Polytechnic
	Telford Rural Polytechnic
	The Open Polytechnic of New Zealand
	Unitec New Zealand
	Universal College of Learning
	Waikato Institute of Technology (Wintec)
	Waiariki Institute of Technology
	Wellington Institute of Technology
	Western Institute of Technology at Taranaki
	Whitireia Community Polytechnic

Universities	Auckland University of Technology
	Lincoln University
	Massey University
	The University of Auckland
	University of Canterbury
	University of Otago
	The University of Waikato
	Victoria University of Wellington

The intention of the exercise was to describe evaluation contexts using publically accessible material and to gain as full a picture as possible about evaluations without actually visiting the institutions. The assumption underlying this approach was that such public information would be information concerning its claims about its evaluation context that an institution has decided is important enough to make available. This material would then provide some baseline material against which the more specific detail about the three institutions participating in the study and the views and perceptions of evaluations gathered through the other data-gathering methods, the questionnaire and the interviews, could be compared and contrasted. Essentially, the expectation was that the material would provide some sense of institutional perspective on student evaluations in the wider New Zealand tertiary education context.

To commence the process, a broad look was taken at the websites in a general scanning exercise. In the first instance, the websites were scanned for evidence in institutional

statements and policies that student evaluations of courses and teaching were carried out in the institution(s). Second, each website was searched for further information about the evaluation/appraisal systems and procedures in place in each institution, including whether consultation about the results of evaluations/appraisals was available. The search provided a way to assess the degree to which the institution explicitly supports the use of evaluation results for whatever purposes it considers important (accountability or development) through the provision of staff and other resources. This initial exploration of websites quickly showed that more specific information about the provision of consultation on the outcomes of evaluation activity was difficult to obtain.

As a result of this general scanning exercise, it was decided that the following aspects would guide and frame a more detailed environmental scan of the institutions:

- a) The *purposes* of student evaluations – to gain a sense of the institution’s reasons for its student evaluation system and processes.
- b) Whether evaluations were *mandated* – to understand expectations about the conduct of student evaluation activity and whether, and possibly how, student evaluation was linked to (a) above.
- c) The *frequency* of evaluations – to find out how often, and by what means, evaluation activity occurred across the institution.
- d) Whether the institution had an *educational development centre* – to gain a sense of whether evaluation was seen as developmental exercise, and whether the purposes identified in (a) could be linked to support in the form of staff and other resources to facilitate any resulting developmental process.
- e) Whether the institution had dedicated *evaluations staff/administrators* – to understand the processes involved in the centralised student evaluation system within the institution.

2.2.3 Results of the scan

Table 2:2 shows a snapshot of the information that was accessible online.

As can be seen from Table 2:2, there is evidence that all eight of the universities and 12 of the polytechnics conduct student evaluations/appraisals of teaching and courses. There is also evidence that 11 of these institutions have specific policies regarding student evaluations and

four of them do not have policies. It was unclear at the time of the scan whether the remaining 13 institutions have policies or not. Fifteen institutions state that they use evaluations to inform teaching and/or course improvement and 12 institutions state that results of evaluations are used for summative purposes such as audits, reviews and performance appraisals.

Six institutions clearly mandate student evaluations of courses and/or teaching. Fifteen institutions in total carried out evaluations on a regular basis, and it could therefore be assumed that evaluations are, to some extent, implicitly mandated at these institutions. All of the universities and seven of the polytechnics have an educational development centre. All of the universities and seven of the polytechnics have an evaluation unit or staff responsible for administering student evaluations.

Table 2:2 Summary of Key Features of Student Evaluation Systems in New Zealand Institutions

	Official documents contain some evidence that the institution conducts student evaluations of courses and/or teaching	Institution has a specific policy or procedure regarding student evaluations		Stated purpose of evaluations		Evaluations are clearly mandated	Evaluations are carried out			Institution has an educational development centre	Institution has an evaluation coordinator/administrator
		yes	no	teaching & course improvement	promotion / appraisal/ audit		at least once a year	biennially	every 3 years		
Polytechnics	12	7	1	8	5	3	7		1	7	6
Universities	8	4	3	7	7	3	1	2	4	8	8
Total Institutions	20	11	4	15	12	6	8	2	5	15	14

A more detailed overview of these results, broken down according to institution, is provided in Table 2:3. Although some of the institutions' websites and official documents had information about whether students have access to the results of the evaluations and whether there is a requirement for action based on data from surveys (and, if so, what requirements), this is not included in Table 2:3, and could be the focus of further investigation as it will inevitably have a bearing on staff engagement with evaluation.

It can be seen from the information presented in Table 2:3 that all institutions' websites surveyed declared some sort of evaluation system, and that there is a spread of approaches, structures and processes, and variation in the policies. Different types of institutions – polytechnics and universities – mean that there are different accountability, quality assurance/enhancement expectations and regimes (see section 2.2.1), but all seem to aim to use their teaching evaluation/appraisal systems to contribute to quality assurance monitoring and development in some way.

The following section provides more information about the three institutions involved in the project and serves to highlight some of the specific detail related to student evaluations/appraisals and the variation in the institutional type. The variation, in accountability and quality assurance/enhancement expectations and regimes, between universities and polytechnics in New Zealand is highlighted and also, how variations exist between institutions of the same type, in this instance, between two universities.

Table 2:3 Details of Environmental Scan, by Institution

Institution	Institutional statements & policies that are available online	Mention of student evaluations occurring is included in available institutional document(s)	Institution has a specific policy &/or procedure regarding student evaluations	Stated purpose of evaluations	Evaluations are mandated	Frequency of evaluations	Institution has an Education Development Centre	Includes evaluation staff
Aoraki Polytechnic	Charter and Student Charter	Yes. In charter and student charter		Teaching and course improvement				
Bay of Plenty Polytechnic	Visioning statement and Quality Management Systems (2009) document	Yes. In QMS document	Yes. Included in QMS document	Teaching and course improvement. Summative, as part of the progression process		Teaching evaluations – at least once a year (permanent staff)	Pikiarero (Teaching and Learning Development Centre) No further detail easily accessible online	
Christchurch Polytechnic Institute of Technology	Academic Policies and Procedures manual		Yes	Teaching and course improvement		All students are given the opportunity to provide formal feedback at least annually	CED (Centre for Educational Development) New centre. No further detail easily accessible online	
Eastern Institute of Technology	Charter and Academic Statutes	Yes					Staff Development Centre, no further detail easily accessible online	

Institution	Institutional statements & policies that are available online	Mention of student evaluations occurring is included in available institutional document(s)	Institution has a specific policy &/or procedure regarding student evaluations	Stated purpose of evaluations	Evaluations are mandated	Frequency of evaluations	Institution has an Education Development Centre	Includes evaluation staff
Manukau Institute of Technology	Charter, Strategic Plan, and Academic Regulations	No	Policies are password protected				Academic Development Centre	An Evaluation Unit. No further detail easily accessible online
Nelson Marlborough Institute of Technology	Academic Statute. Teaching and Learning Statute. Student Surveys Statute	Yes, in Teaching and Learning Statute and Student Surveys Statutes	Student Surveys Statute	Improvement. Evidence of improvement. Meet TEC reporting requirements		Carried out at the discretion of the Quality Committee		There is a survey coordinator. No further detail easily accessible online
NorthTec	Academic Regulations and Statutes	No						
Otago Polytechnic	Charter	Yes	Yes	Improvement. Inform course and programme reviews	All staff are required to seek feedback on at least an annual basis	Courses are evaluated at least once a year	Organisational Research (part of Academic Services) and EDC	Organisational Research Officer. Responsible for administering/ coordinating surveys
Southern Institute of Technology	No							

Institution	Institutional statements & policies that are available online	Mention of student evaluations occurring is included in available institutional document(s)	Institution has a specific policy &/or procedure regarding student evaluations	Stated purpose of evaluations	Evaluations are mandated	Frequency of evaluations	Institution has an Education Development Centre	Includes evaluation staff
Tai Poutini Polytechnic	Mission Statement. Strategic plan. QMS document	In QMS document	Referred to in performance appraisal policy, which includes procedure for student evaluation of tutor	Inform performance appraisal and professional development activity		Annually		Teaching Development Coordinator
Tairawhiti Polytechnic	Strategic Plan	Student satisfaction is noted as a good teaching indicator. Survey process to be developed	No					
Telford Rural Polytechnic	Investment Plan. QMS policies	Yes	Yes	Staff and programme improvement. Appraisal	Yes	Full-time staff twice a year, part-time staff once a year. Full year programmes must include a mid-year and exit survey.		

Institution	Institutional statements & policies that are available online	Mention of student evaluations occurring is included in available institutional document(s)	Institution has a specific policy &/or procedure regarding student evaluations	Stated purpose of evaluations	Evaluations are mandated	Frequency of evaluations	Institution has an Education Development Centre	Includes evaluation staff
Open Polytechnic	Investment Plan Annual report	No					Centre for Academic Development. No further detail available online	
Unitec	Learning Quality Plan	Yes	Yes	Improvement and development of courses, programmes and teaching	There is a rolling triennial schedule of mandatory course evaluation	Standard teacher evaluations administered according to an annual plan	Could not find but indications that they have one	Planning Services responsible for preparing surveys and reports
Universal College of Learning	Annual Report. Investment Plan	No		Policies not available online-unclear whether there is one			Educational Delivery and Innovation	EDI is Responsible for administrating and reporting on student surveys
Wairiki Institute of Technology	Academic Statute. Charter	Yes, in Statute						
Waikato Institute of Technology (Wintec)	Annual Report	No						
Wellington Institute of Technology	Charter. Annual Report	No						

Institution	Institutional statements & policies that are available online	Mention of student evaluations occurring is included in available institutional document(s)	Institution has a specific policy &/or procedure regarding student evaluations	Stated purpose of evaluations	Evaluations are mandated	Frequency of evaluations	Institution has an Education Development Centre	Includes evaluation staff
Western Institute of Technology at Taranaki								
Whitireia Community Polytechnic	Charter. Annual Report	Yes				All teaching staff and support services are evaluated by students each year (p.8, Charter)		
The University of Auckland	Charter	No	Yes. Policy and a handbook on the evaluation of courses and teaching	Monitoring and reporting on the quality of courses and teaching. Quality enhancement		Once every three years. Rolling three-year evaluation plan	Centre for Academic Development. Advises on pedagogical issues and professional development needs arising from evaluations	Evaluations and Scanning Centre processes student evaluation forms and reports and results. Quality office runs student evaluations website (password protected)

Institution	Institutional statements & policies that are available online	Mention of student evaluations occurring is included in available institutional document(s)	Institution has a specific policy &/or procedure regarding student evaluations	Stated purpose of evaluations	Evaluations are mandated	Frequency of evaluations	Institution has an Education Development Centre	Includes evaluation staff
Auckland University of Technology	Learning and Teaching Framework	Yes	Policies are password protected. Has a procedure	Inform paper and programme reviews. Improvement of courses and teaching		Every three years	Yes (CPED)	Has an Institutional Research Unit. Responsible for administering SEP (Student Evaluation of Papers)
The University of Waikato	Charter	No	Yes	Ongoing professional development. Improvement of Teaching	Yes	Every two years	Teaching Development Unit (TDU)	Teaching Developer (Evaluation & Quality)

Institution	Institutional statements & policies that are available online	Mention of student evaluations occurring is included in available institutional document(s)	Institution has a specific policy &/or procedure regarding student evaluations	Stated purpose of evaluations	Evaluations are mandated	Frequency of evaluations	Institution has an Education Development Centre	Includes evaluation staff
Massey University	Charter	Yes	No	The Massey Online Survey Tool replaces the previous SECAT (Student Evaluation of Content, Administration & Teaching). The main purpose of the Course Survey is for accountability, quality assurance and continuous improvement and the main purpose for the Teacher Survey is teacher development and improvement		The Massey Online Survey Tool Course Survey to evaluate courses is compulsory at least every third offering with the time between evaluations not exceeding two years. Teacher Evaluation is not compulsory	Until 2011, the Centre for Academic Development and eLearning (CADEL) coordinated the SECAT process. From 2012, Centres for Teaching and Learning will be established. The Student Engagement and Evaluation Unit (SEEU) is part of the National Centre for Teaching and Learning	Student Engagement and Evaluation Unit (SEEU) Manager and three staff
Victoria University of Wellington	Teaching and Learning Strategy. [Password protected]		No				University Teaching Development Centre (UTDC)	Evaluations Administrator

Institution	Institutional statements & policies that are available online	Mention of student evaluations occurring is included in available institutional document(s)	Institution has a specific policy &/or procedure regarding student evaluations	Stated purpose of evaluations	Evaluations are mandated	Frequency of evaluations	Institution has an Education Development Centre	Includes evaluation staff
University of Canterbury	University Plan- Teaching and Learning Plan 2011-2013	Review courses and/or programmes for relevance of content and of pedagogical approaches, seeking feedback from students and stakeholders where appropriate		Yes. Student Surveys are intended to allow students to reflect on their learning and offer a regular opportunity to comment on their courses and to provide their perceptions of the teaching involved	Yes, if low scores when last surveyed. For promotion.	Every three years	Academic Development Group (ADG)	Student evaluation surveys are processed by the Centre for Evaluation and Monitoring (CEM). Survey Administrator
Lincoln University	Investment Plan. Teaching and Learning Plan	Section in Teaching and Learning Plan		Promotion and teaching improvement	Subject evaluations are mandated. Teacher evaluations are voluntary	Subject evaluations are carried out biennially	Teaching and Learning Services. A student evaluations section	Administrator of Teaching Evaluations
University of Otago	Charter. Academic Promotions Policy	Otago Teaching Profile guide. Academic Promotions Policy	No policy. Otago Teaching Profile guide	Promotion. Improvement of teaching and courses		All significant teaching responsibilities should be evaluated at least once during the previous three years	Higher Education Development Centre (HEDC).	Three evaluation staff

2.3 Setting - The Three Institutions Involved in the Project

Section 2.2 provided information about the New Zealand tertiary environment with particular reference to the public statements about student evaluation made by universities and polytechnics (in section 2.2.2). This material provides the backdrop against which the current study was set.

In the following sections, each institution that participated in the study – the University of Otago (OU), The University of Waikato (WU) and the Otago Polytechnic (OP) – is described. First, for each institution, a general description is given including a brief overview of its history and its current vision, mission and values, and a short description of governance. Following this, information about teaching and teaching development at the institution is presented. Included are significant features of the institution's systems and processes that have a bearing upon student evaluations/appraisals. Reference is also made to the summaries presented in Table 2:4 showing the vision, mission and values of the three institutions involved in the study, and Table 2:5 outlining the contexts around the use of evaluation/appraisal at each institution.

2.3.1.1 University of Otago

2.3.1.1.1 Background

The University of Otago was established in 1869 and opened in 1871. In 1874 the University became an affiliate of the federal University of New Zealand, which was established in 1870, meaning that degrees were conferred by the federal body not by the University of Otago itself. In 1961 when the University of New Zealand was disestablished, the power to confer degrees was transferred back to the University of Otago. The University thus has a long tradition and its operation and structure today reflects that history.

Today, the overarching body governing the University is the Council, presided over by the Chancellor. This body receives advice from heads of departments. The Vice-Chancellor, similar to a chief executive officer found in many organisations,

Table 2:4 Vision, Mission and Values of Each Institution Involved in the Project

	OU (University of Otago, 2003, pp. 1, 3-4)	WU (The University of Waikato, 2009, p. 2)	OP (Otago Polytechnic, 2011, p. 1)
Vision	A research-led University with an international reputation for excellence	<p>The <i>Vision</i> for The University of Waikato, informed by our existing high-quality teaching and research, is as follows:</p> <p>We will</p> <ul style="list-style-type: none"> • deliver a world-class education and research portfolio • provide a full and dynamic university experience which is distinctive in character • pursue strong international linkages to advance knowledge 	Through our innovation in education and the outstanding experiences our learners enjoy, we will be recognised as New Zealand’s leading polytechnic
Mission	The University of Otago will advance, preserve and promote knowledge, critical thinking and intellectual independence to enhance the understanding, development and well-being of individuals and society. It will achieve this by building on foundations of broad research and teaching capabilities, unique campus learning environments, its nationwide presence and mana, and international links	To combine the creation of knowledge through research, scholarship and creative works with the dissemination of knowledge through teaching, publication and performance	<p>We inspire learning as we develop capable practitioners for Otago and New Zealand.</p> <ul style="list-style-type: none"> • Our graduates are work ready and prepared for participation in a global workforce • People from all backgrounds and prior education achievements have opportunities to develop further at Otago Polytechnic • We specialise in experiential learning at all levels from foundation skills to advanced professional practice, offering innovative programmes of quality and relevance • Our communities influence what we do and how we do it, benefiting from our

	OU (University of Otago, 2003, pp. 1, 3-4)	WU (The University of Waikato, 2009, p. 2)	OP (Otago Polytechnic, 2011, p. 1)
			joint endeavours <ul style="list-style-type: none"> We are socially responsible, committed to deploying our resources in the interests of our immediate communities and a sustainable future
Values	In pursuing its mission, the University of Otago adheres to the following core values: <ol style="list-style-type: none"> Intellectual Independence & Academic Freedom Excellence Partnership Leadership Collegiality & Collaboration Knowledge Equity & Ethical Standards Consultation Stewardship 	The University of Waikato places a high value on: <ul style="list-style-type: none"> Partnership with Māori Acting with integrity Celebrating diversity Promoting creativity 	In carrying out our mission we will honour the values of: <ul style="list-style-type: none"> Caring: we are respectful, generous, welcoming and inclusive Responsibility: we do the right thing, we are accountable for meeting our commitments Partnerships: we strive to work and learn together Learning: we seek to learn from all that we do Excellence: we seek always to do our best; we set high expectations

Table 2:5 The Institutional Context Around Evaluation/Appraisal of the Three Participating Institutions

Note: Unlike OP and OU, WU uses the term ‘appraisal’ to mean the summative questionnaire used at the end of a semester, while ‘evaluation’ is specifically used to describe formative types of data collected during the semester by teachers to inform and develop their teaching. Throughout this report, we use ‘evaluation/appraisal’ to signify this variation in the use of terminology.

	OU	WU	OP
Evaluation/appraisal institutional context			
Who in each institution has responsibility for teaching development?	Higher Education Development Centre (HEDC)	Teaching Development Unit (TDU) and the Waikato Centre for eLearning (WCeL)	General Manager Academic Services who delegates that responsibility to the Educational Development Centre
Mandatory teacher training for new staff	No	No	Yes. All new staff are required to obtain the Graduate Certificate in Tertiary Learning and Teaching unless they already hold a tertiary teaching qualification
Probation	Five-year Confirmation Path for academic staff	Not required	Not required
Ongoing appraisal/review	Regular reviews over a 5-year period form part of the Confirmation Path process. Once confirmed, annual Performance Appraisal	Annually via Professional Goal Setting (PGS) process	No
Mandatory teaching development	No	No	Yes. All staff are required to attend mandatory staff development days
Requirements for professional tertiary teaching qualification	No	No	Yes. All lecturing staff are expected to have obtained the Graduate Certificate in Tertiary Learning and Teaching within 3 years of their appointment
Requirements for data around teaching in	Teaching evaluations mandatory, course evaluations optional – expectation that	Only paper and teaching appraisal results are required	Portfolio of evidence including Annual Student Feedback on Teaching Report

	OU	WU	OP
promotions process	results are discussed as part of the Otago Teaching Profile		and Peer Feedback report from at least 5 colleagues
Recognition for teaching	Through local and national teaching awards. Annual institutional award	Through local and national teaching awards	Through local and national teaching awards and Institution annual excellence awards
Evaluation/appraisal process			
How often and when to evaluate/appraise	<p>No stipulation on when evaluations are run.</p> <p>Recommendation that Course Evaluation Questionnaires be regularly and routinely used by all departments as needs arise, but at least on a three-year cycle. Departments should establish internal processes to monitor this requirement, which should be part of regular and routine departmental-based evaluation policy</p> <p>Teaching evaluations are optional. Formative informal evaluations can be run at any time</p>	<p>Once every two years for paper and teaching</p> <p>Formative informal evaluations can be run at any time after week 4, but formal appraisals are mainly at mid or end of semester</p>	<p>Determined by Head of School, Programme Manager or individual lecturer, but mainly take place within last 4–6 weeks of a programme or course</p>
Flexibility with evaluation/appraisal medium	Can be either paper-based or online evaluation	Can be either paper-based or online appraisal	Can be either paper-based or online evaluation
Flexibility with questions	<p>Teaching evaluations: 5 mandatory core questions. Staff have to choose a further 5 questions from a catalogue of 45 questions. The questions are 5-rating Likert-scale with a descriptor at each end. General comments are permitted on back of questionnaire</p> <p>Course evaluations: Up to 30 questions chosen from a catalogue of 289 general</p>	<p>Have 8 core questions for paper and teaching appraisals, with a 5-point behavioural observation Likert scale. There are also two formative questions to gather student comments. Staff can choose further summative or formative questions from an item bank, or develop their own questions in conjunction with TDU for specific situations</p>	<p>Fixed questions for student feedback on both Teaching and Course evaluation</p> <p>Student feedback on teaching: 10 rating questions and 2 sections for comments</p> <p>Course evaluation form: 11 rating questions and 2 sections for comments. Lecturers can request additional questions for the course evaluation in</p>

	OU	WU	OP
	and clinical questions and can be customised to suit the particular course. The questions are 5-rating Likert-scale with a descriptor at each end. Also permitted to have open-ended comments if wanted		special situations
Return of evaluation/appraisal results	<p>Evaluation data returned as soon as possible. For paper evaluations the originals of questionnaires with unedited student comments also returned. Online evaluations have the comments included in the report for course evaluations and as a separate report for the teacher evaluations.</p> <p>Teacher evaluation results are provided in a Summary report as distribution percentages. All the evaluations run within a particular year are shown on the report and there is no choice to exclude a set of results.</p> <p>Course evaluation results are provided in a detailed report showing distribution data and percentages. A single report is created for each evaluation run.</p>	<p>Summative appraisal data returned as soon as possible. Formative open question data returned after the marks have been finalised. Unedited student comments are also included. Online appraisals also sent back with open questions included in reports.</p> <p>Annual reports of paper appraisal data are sent to each department, Faculty and ultimately the Teaching Quality Committee. These reports provide each level with benchmarks, which relate directly to the KPIs in the <i>Teaching and Learning Plan 2010 – 2012</i>.</p>	<p>Student Feedback on Teaching Report returned to individual lecturer with tables and graphs of the ratings provided by students and a summary of the unedited comments made by students.</p> <p>Course evaluation report has tables and graphs of the ratings questions and summary of unedited comments. The report is returned to the requesting lecturer and/or programme manager.</p> <p>Edited copies (information identifying individual lecturers is removed) are placed on the institution's performance portal where all key reports are held and are available to all staff.</p> <p>Where ratings to any questions are less than 80% agreement a follow-up action report is sent to the Head of School who is required to report to the General Manager Academic Services on the School's response to the ratings. Expectation that School's response will also be provided to students who provided the feedback.</p>
Requirement to feed back to	Not mandatory. Good practice of	To next cohort via paper outlines only.	Not compulsory but encouraged.

	OU	WU	OP
students	feeding back to students is promoted and encouraged.		
Reporting of evaluation/appraisal data	<p>Teaching evaluation results are returned only to the requesting teacher, but the teacher reports their teaching evaluations to Heads of Departments and promotions committees via Confirmation, Annual Performance Appraisal and Promotions processes.</p> <p>Course evaluations are returned to course coordinator and copied to the Head of Department.</p>	<p>Reporting of paper and teacher appraisal to individuals.</p> <p>Aggregated paper appraisal data sent to Chairs of Departments (CODs), Deans and Teaching Quality Committee.</p> <p>Teacher appraisal data reported CODs by individual via Professional Goal Setting (PGS).</p>	<p>Student Feedback on Teaching Report is returned to the requesting lecturer and is treated as private information for the lecturer. That lecturer may use the information as part of their annual performance appraisal meeting or when applying for a salary review.</p> <p>A Head of School may request a student feedback on teaching survey and report where they have concerns about an aspect of that lecturer's teaching.</p> <p>All course evaluation reports are placed on the performance portal.</p> <p>From 2011 aggregated data from all course evaluations by school and course will also be provided on performance portal and on student portal.</p>

convenes the Senate which takes advice from various committees and groups, including the four Divisions: Commerce (School of Business), Science, Humanities and Health Sciences. Each Division is headed by a Pro-Vice-Chancellor. Alongside the Vice-Chancellor and the four Pro Vice-Chancellors, there are two Deputy Vice-Chancellors, having oversight of Academic and International and Research matters across the University. In addition, there is a fifth Pro Vice-Chancellor, focussing on International matters as well as a Chief Operating Officer. Nine non-teaching Divisions complete the structural organisation of the University. These include: Academic, Research, Human Resources, Accommodation Services, Financial Services, Information Technology Services, Marketing and Communications, Property Services, and Student Services (University of Otago website <http://www.otago.ac.nz/about/history/index.html>).

The University of Otago has its main campus in Dunedin with campuses in Christchurch, Wellington and Invercargill and an information and teaching centre in central Auckland. In 2010, there were 21,139 individual students (19,918 EFTS) enrolled across 194 undergraduate and postgraduate degree, diploma and certificate programmes (University of Otago, 2011). In 2010, there were 1,167 (FTE) staff involved in both teaching and research employed across the institution (University of Otago, 2011).

2.3.1.1.2 Teaching

Oversight of teaching and research is the role of Senate. The Committee for the Advancement of Learning and Teaching (CALT), convened by the Deputy Vice-Chancellor (Academic and International), is a committee of Senate and has the role of supporting and promoting excellence in teaching and learning across the University. Included in the activities of the committee is the development and promotion of guidelines for teaching, namely *The Teaching and Learning Plan*. CALT also acts as the advisory body for the Higher Education Development Centre (HEDC), which has the role of supporting the development of teaching and learning across the University. (See further description of the role of HEDC in section 2.3.1.1.3). CALT does not have an auditing role; the quality of programmes and courses being the responsibility of two other committees of Senate, namely, the Board of Undergraduate Studies (BUGS) and the Board of Graduate Studies (BOGS).

The Teaching and Learning Plan 2005-2010

The current Teaching and Learning Plan 2005-2010 is an update of the 2002 version, which was “developed by a group of academic staff experienced and active in teaching” (University of Otago, 2005, p. 3). The *Teaching and Learning Plan 2010-2012* (University of Otago, 2005) includes within it a series of nine objectives linked to the University’s Profile 2004-2006. Each objective is supported by a series of strategy statements, followed by key performance indicators. The *Teaching and Learning Plan* applies to all levels of undergraduate and postgraduate teaching in the University and is used by Divisions to guide the planning of teaching and learning activities. Within the *Teaching and Learning Plan* explicit mention of the evaluation of the quality of teaching in programmes appears within the list of suggested teaching strategies to support Objective 4: To promote, encourage and support excellence in teaching.

4c: Provide a comprehensive system for evaluating the quality of teaching in programmes, courses and by individual teachers which emphasises development and fosters a desire to improve. (p. 8)

There is no other mention of evaluation in the *Plan*.

A new *Teaching and Learning Plan* that will focus on teacher and student participation was under development during 2011. It will contain a more expansive section addressing student evaluation.

2.3.1.1.3 Teaching development

The Higher Education Development Centre (HEDC) is a central unit within the University and its role is “to work in partnership with staff and students of the University to promote, support and enhance the ideals, knowledge and values of higher education” (HEDC website <http://hedc.otago.ac.nz/hedc/home/About-Us.html>). As such, the unit provides academic staff development support and student learning development support across the institution. HEDC coordinates and teaches a Postgraduate Certificate and Diploma in Higher Education and supervises Masters and PhD students. HEDC is also the body that administers the University’s student evaluation of course and teaching systems.

Unlike the University of Waikato, the University of Otago does not have a formal policy on the evaluation of teaching and courses. Rather, imperatives for staff to be involved in evaluation come from (a) advice about good teaching, learning and curriculum (promoted through CALT, HEDC and other units such as the Distance Learning Office and their associated activities; through the quality assurance roles played by BOGS and BUGS; through teaching and learning committees and groups active within departments and Divisions), and (b) Human Resource policies on confirmation, promotion and annual performance.

The compulsory development of an *Otago Teaching Profile* (<http://hedc.otago.ac.nz/hedc/etc/Otago-Teaching-Profile.html>) by academic staff involved in teaching is one way that encourages thinking about the links between the accountability and the developmental uses of student evaluation data. The document has to be included as part of confirmation and promotion submissions. The core element of this document is a staff member's personal statement about teaching, including his or her philosophy about teaching and learning in context, the teacher's goals and achievements, all supported by evidence to demonstrate success and effectiveness. Reference to student evaluation of teaching data is compulsory, but student evaluation of course data may also be included. In fact, encouragement is given to teachers to include reference to a variety of other data sources also, so that a fuller picture of their teaching context and situation can be communicated. In this way, the assumption is that the quality of teaching cannot be concluded using only one kind of data. However, only the inclusion of the student evaluation of teaching data is compulsory.

The *Otago Teaching Profile* is intended to be an evolving document that staff contribute to over time. The spirit of the document is such that the inherent questions and challenges that arise in any teaching are recognised, so that whatever the results of student evaluations and other evaluative data, the expectation is that the staff member discusses his or her effectiveness, drawing on the data in a realistic open way and makes explicit the steps he or she has put in place to address any issues.

2.3.1.1.4 Other features of significance for evaluation/appraisal

Confirmation Processes at the University of Otago

When an academic staff member first attains a position at the University he or she will enter a five-year probationary period called ‘confirmation’. For each staff member on Confirmation Path a set of goals is negotiated around:

- teaching, assessment and curriculum development
- establishing a research profile and being an independent researcher
- service contributions appropriate to the new staff member’s appointment level.

The staff member’s success and effectiveness in their role is gauged around how well their goals have been achieved. Reports on progress towards achievement of the goals are made by the staff member each year for the period of confirmation, with movement to a permanent position attainable any time between three and five years, depending on whether the set goals have been achieved. Part of reporting on achievement of teaching, assessment and curriculum development goals is the mandatory submission of an *Otago Teaching Profile*.

Details about the confirmation process at the University of Otago are available at <https://docushare.otago.ac.nz/docushare/dsweb/Get/Document-38753/ConfirmationOfAppointmentPolicySep10.pdf>.

Promotion Processes at the University of Otago

“Promotions generally refer to cases where the job title changes and progression refers to a situation where someone moves up the salary scale but retains the same job title” (University of Otago Human Resources, 2011, p. 4). When a staff member is ready, he or she may apply for promotion. In the application, which must outline the case made to demonstrate competence (sustained to sustained high level competence), evidence must be provided around the following three areas:

- teaching, assessment and curriculum development (except for research only roles)
- research, scholarship, professional practice or performance activities (except for teaching only roles)
- service to the University and the community and demonstration of collegiality within the Department/School.

(University of Otago Human Resources, 2011, p. 11)

Once again, the *Otago Teaching Profile* forms the core statement that must be supplied about the teaching, assessment and curriculum development area.

Details about the promotion process at the University of Otago are available at <https://docushare.otago.ac.nz/docushare/dsweb/Get/Document-11235/AcademicPromotionsPolicy.pdf>. The document contains lists of prompts that suggest examples of activity for which applicants might provide evidence of sustained to sustained high level competence.

2.3.1.2 The University of Waikato

2.3.1.2.1 Background

After a long fight by the University of South Auckland Society (founded in 1956) The University of Waikato became a reality in 1964. Opened by the then Governor-General, Sir Bernard Fergusson, the University was mainly farmland with a few temporary buildings and a handful of staff. Today The University of Waikato has its main campus in Hamilton with a satellite campus, in partnership with the Bay of Plenty Polytechnic, at Tauranga. There are currently 14,000 students enrolled across 80 qualifications ranging from Foundation programmes to PhDs. The University has seven Faculties/Schools including Management; Education; Science and Engineering; Computing and Mathematical Sciences; Law; Arts and Social Sciences; and Māori and Pacific Development. Currently there are 633 (FTE) academic staff employed within these Faculties/Schools.

The mission statement of the University of Waikato is:

To combine the creation of knowledge through research, scholarship and creative works with the dissemination of knowledge through teaching, publication and performance.

How the University achieves this mission is the responsibility of Council, which has responsibility for the governance of the University. Chaired by the Chancellor, it has 18 members, including representatives of the Minister of Education, the Vice-Chancellor, other nominated representatives of the University and external stakeholders.

To advise and support the Vice-Chancellor around management of the University, there is a Senior Leadership Team (SLT). This team meets monthly to consider and address matters of

strategic and operational importance as described in the University Strategy and suite of Plans. For example:

The University of Waikato is committed to providing a world-class, relevant and sustainable programme of teaching and learning (University Strategy 2010-2013 - Goal 1).

The SLT consists of 13 key academic leaders including the Deputy Vice-Chancellor, the Deans of each of the seven faculties mentioned above, and five Pro Vice-Chancellors, with portfolios that cover: International; Māori; Postgraduate; Research; and Teaching and Learning. Also on the SLT are two Assistant Vice-Chancellors, Executive and Operations, and the Heads of Divisions of Finance; Facilities Management; Human Resource Management; Information Technology Services; and Student and Academic Services.

2.3.1.2.2 Teaching

Oversight of teaching within the University of Waikato is through the Teaching Quality Committee, a standing committee of Academic Board. Chaired by the Pro Vice-Chancellor Teaching and Learning, the committee has ownership of a number of key strategic documents around teaching and learning including *The Teaching and Learning Framework* and the *Teaching and Learning Plan 2010-2012*

(<http://www.waikato.ac.nz/about/corporate/tlmgplan.shtml>).

The *Teaching and Learning Framework 2010-2012* was written by seven different groups, each selected for their expertise around a particular goal within the plan. This approach to developing the plan stimulated far-ranging conversations about paper and teaching quality, continuous improvement, student support, supervision, e-learning and facilities at all levels of the University. Distilled from the *Teaching and Learning Framework*, the *Teaching and Learning Plan 2010-2012* includes within it performance indicators and benchmarks to assess not only teaching quality and the professional development of staff, but also student learning development and support activities. The use of key performance indicators and benchmarks allows a comparison of similar areas, either internal or external to the institution, with a view to achieving best possible practice in that particular area.

Policy on the Evaluation of Teaching and Papers

Embedded within the *Teaching and Learning Plan* is the *Policy on the Evaluation of Teaching and Papers*. The policy highlights the requirement for both formative and summative data for teaching development and institutional processes. Concurrently, a set of guidelines was developed to support the policy and includes sections on the use of both appraisal and evaluation data (http://www.waikato.ac.nz/tdu/pdf/booklets/15_Appraisals.pdf).

Other related policies and guidelines that have an impact on teaching include the *Academic Workloads Policy* (<http://www.waikato.ac.nz/official-info/index/docs/academic-workloads-policy>), which divides academic time into teaching, research and administration on a 40:40:20 split, respectively. Alongside this policy is the *Teaching Buy-out Policy*, which allows academic staff to ‘buy-out’ teaching time to fulfil research obligations. Another pertinent policy is the *Paper Outline Policy* and associated template (<http://www.waikato.ac.nz/official-info/index/docs/paper-outline-policy>). This policy outlines the requirements of a paper, including the need to mention student feedback and lecturer responses from the last appraisal for that particular paper. This information includes any changes to the paper or teaching that have occurred as a consequence of the student feedback.

2.3.1.2.3 Teaching development

There are two main units charged with responsibility for the development of academic teachers at The University of Waikato; the Teaching Development Unit (TDU) and the Waikato Centre for eLearning (WCeL). These two separate units are hosted in the Faculty of Education. Both units work collaboratively to provide a variety of generic and discipline-specific development opportunities. Both units work closely with the Pro-Vice Chancellor (Teaching and Learning), and the Teaching Quality Committee to ensure consistency between institutional and individual teacher professional development needs.

The TDU aims to “help staff to develop the competencies, skills and pedagogical knowledge that they need to provide high quality teaching”. This is achieved through a range of different initiatives including: a workshop series (run twice a year); one-to-one or departmental consultancies; teaching advocacy at Faculty level; appraisals and evaluations of teaching and papers; and resources to support all aspects of tertiary teaching. The TDU also runs the Postgraduate Certificate in Tertiary Teaching, a non-compulsory, practice-based postgraduate

tertiary teaching qualification. In addition, the TDU oversees the University's local Teaching Awards and helps staff with portfolio development for the National Tertiary Teaching Awards.

WCeL is charged with supporting individuals, departments and the University, around technology to support pedagogy.

2.3.1.2.4 Other features of significance to evaluation/appraisal

Use of appraisals in promotions processes

Currently, promotion processes require a summary of appraisal data. These data are provided via the Academic Staff Portfolio (ASP), an online repository for information required for submission by individual academics as part of the University's promotion processes. The ASP is auto-populated by appraisal information provided by the TDU. The appraisal data from the TDU are the only data used in the promotion process to assess the quality of a paper and the teaching.

Professional Goal Setting (PGS)

Each year general and academic staff members are required to attend a PGS meeting to determine requirements for work-related issues, individual performance and future plans. PGS results in an individual Work Plan, which will include goals for the following one to three years, consistent with University Strategy and Plans, an individual professional development programme, and, for academic staff, individual research, teaching and community service plans, as appropriate. For this process, academic staff are expected to provide appraisal results for their teaching so that appropriate professional development opportunities are provided. Further information around the PGS process can be found at: <http://www.waikato.ac.nz/hrm/pgs/>.

2.3.1.3 The Otago Polytechnic

2.3.1.3.1 Background

Otago Polytechnic traces its ancestry back to the Dunedin Technical School, which was established in 1889 to provide evening classes for working people. In 1909 it expanded to

offer day classes for secondary school pupils. In 1914 the name was changed to the King Edward Technical College.

In 1921 the college took over the Dunedin School of Art, which had been established in 1870. The college expanded further by taking on the evening and day time education of apprentices, technicians and professionals. In 1966 the college was split into a secondary school (later re-named Logan Park High School) and Otago Polytechnic, which opened on 1 February 1966.

Otago Polytechnic has its main campus in Dunedin with a smaller campus in Cromwell and outreach sites at Queenstown and Wanaka. Programmes are also run at the Waikato Institute of Technology (Wintec) in Hamilton and the Bay of Plenty Polytechnic in Tauranga. It also carries out distance-based learning in subject areas ranging from Veterinary Nursing to Midwifery.

There are 447 full-time equivalent staff (FTE) employed and 3680 equivalent full-time students (EFTS). In 2010, 8106 individual students were enrolled. Otago Polytechnic offers in excess of 165 qualifications ranging from Level 2 certificates to Masters programmes. Forty per cent of all qualifications offered are at degree level or higher.

The Otago Polytechnic is a crown/public entity governed by its own Council, which comprises eight members. Four members, including the chair and deputy chair, are appointed by the Minister of Tertiary Education, with four appointed by the Council itself. The Council has four subcommittees: Finance and Audit; Komiti Kawanataka; Staff; and Student. The Chief Executive reports to the Council and works with the Leadership Team. Each academic area and service area has ongoing self-assessment that is evaluated and reviewed by the Leadership Team. Group Managers who oversee many Schools also report to the Chief Executive and the Leadership Team, and there are several Heads of Schools and Heads of Service Departments.

2.3.1.3.2 Teaching

Oversight of teaching at Otago Polytechnic is through the Academic Board, chaired by Chief Executive Officer or nominee. The Academic Board's roles and responsibilities include the teaching, learning, research and technology transfer strategies and activities of the Polytechnic, the Polytechnic's quality systems and processes, academic, management and

council policies. The Academic Board is also responsible for academic performance of the Polytechnic, and of the programmes offered by the Polytechnic. The Board has three standing committees: the Research and Postgraduate Committee; the Quality and Approvals Committee; and the Teaching and Learning Committee.

The Teaching and Learning Committee leads the development of the strategic focus and direction of teaching and learning, maintains the institution's teaching and learning strategic framework, provides advice on staff development related to the teaching and learning strategy and annually reviews the effectiveness of teaching and learning throughout the institution. The key strategic documents and policies around teaching and learning include *Refocussing on Teaching and Learning: Lifting our Game* (<http://enhancingteachingandlearning.blogspot.com/2008/03/refocussing-on-teaching-and-learning.html>) and the policies on *Self- Assessment and Internal Evaluation, Academic Board* (http://www.otagopolytechnic.ac.nz/fileadmin/DepartmentalResources/Marketing/Policies/Academic/AP0101.07_Academic_Board.pdf) and *Student Surveys and Course Evaluations* (http://www.otagopolytechnic.ac.nz/fileadmin/DepartmentalResources/Marketing/Policies/Academic/AP0700.06_Student_Surveys_and_Course_Evaluations.pdf). The *Self-Assessment and Internal Evaluation* policy ensures the Polytechnic is maintaining and enhancing its quality practices in line with New Zealand Qualifications Authority (NZQA) Self-Assessment and External Evaluations and Review (<http://www.nzqa.govt.nz/providers-partners/registration-and-accreditation/self-assessment/>).

Refocussing on Teaching and Learning: Lifting our Game emphasises putting teaching and learning first. This document consistently applies the current policies and initiatives for quality teaching and learning, reprioritising of resources, including the roles of some staff, ensuring good practices are consistent across the institution, building a stronger culture in support of quality teaching and learning and augmenting the support for teaching and learning. The intended outcomes of the strategy are to improve the measurement of academic performance, act promptly and decisively on the evidence, further lift the capability of teachers, require higher levels of accountability and ensure compliance by teachers with the requirements for student and peer feedback.

The annual School and Programmes review process is a key action of Otago Polytechnic's quality practices that are intended to promulgate good practice and promote effective quality

management throughout the institution by way of self-assessment and review. Otago Polytechnic's self-review and continuous improvements process is focussed on the Institution's strategic priorities of teaching and learning with an emphasis on learner achievement, the quality of the learner experience and on the quality of the work environment. Data from a number of sources including student feedback on course and teaching is used to inform the review process.

The *Student Surveys and Course Evaluations* policy states that all lecturing staff members at the institution are expected to gather student feedback on their teaching and every School is expected to have a plan to gather course evaluation data from students on the courses that make up each qualification. The policy contains several guidelines for receiving student feedback on courses, lecturers and programmes. All lecturing staff at the institution have a professional responsibility to gather students' feedback on their teaching as this is regarded as part of their ongoing consideration of the effectiveness of their workplace. Lecturers are advised to undertake student evaluations on an annual basis. A maximum number of two evaluations for a course is expected. Unless the course receives excellent student feedback and there are no major changes, the Head of Schools/Programmes may decide to run formal evaluation surveys more frequently.

2.3.1.3.3 Teaching development

It is expected that unless they already hold a tertiary teaching qualification, all academic staff commencing with Otago Polytechnic are to engage in a tertiary teaching qualification. They must work towards gaining the Graduate Certificate in Tertiary Learning and Teaching, offered by Otago Polytechnic, within a three-year period, or a Postgraduate Certificate/Diploma in Higher Education from the University of Otago. The academic staff member's workload is adjusted to compensate for the time commitment to gain the qualification, having 80 per cent towards teaching hours and 20 per cent towards study. It is the responsibility of the Head of School to manage the lecturers' workloads and provide the necessary support in conjunction with the Educational Development Centre of Otago Polytechnic who manages and delivers the Graduate Certificate in Tertiary Learning and Teaching programme.

The Graduate Certificate in Tertiary Learning and Teaching programme's overall intention is to support transformative developments within learning and teaching contexts for all teaching staff. It aims to support and assist staff to create, deliver and assess quality learning opportunities offered through a blend of online and face-to-face teaching methods. It is designed to support and extend those working as educators with adults.

The Education Development Unit teaches the Graduate Certificate in Tertiary Learning and Teaching programme. In addition, the Unit teaches specific modules for the National Certificate in Adult Education and Training programme, and offers a Certificate in Mata a Ao Māori. They provide one-to-one consultancy, as needed for staff, in terms of staff development, and offer individual mentoring that arises from lecturers' requests or from the Heads of School.

2.3.1.3.4 Other features of significance to evaluation/appraisal

When a staff member is ready, he or she may apply for promotion or a salary review. The Head of School assists the individual to examine their portfolio, which incorporates any new qualifications, research output, and student and course evaluation results. The individual staff member is responsible for collating the information that includes a letter of application, a letter from the Head of Department of School/Service Manager supporting the application, a Performance Review (no more than nine months old), colleague feedback survey (within the last four months), customer feedback survey (student evaluation data) on courses and teaching. This Performance Review portfolio will then go to the Salary Review Committee.

The Otago Polytechnic Staff Development policy

(http://www.otagopolytechnic.ac.nz/fileadmin/DepartmentalResources/Marketing/Policies/M0461.06_Staff_Development.pdf) also requires the academic to write an individual

Development Plan (IDP), which is a statement that is agreed annually between a staff member and manager and sets out the directions for a staff member's development. The Performance Review, which is written by the Head of School/Manager, incorporates the outcomes and actions that are put into the IDP.

2.4 Summary of Chapter 2 – A Conceptual Framework

This chapter presented an overview of current literature that has addressed key issues related to student evaluation/appraisals, as well as an environmental scan of the New Zealand tertiary scene. A particular focus was on student evaluations/appraisals systems at universities and polytechnics. Beginning with an overview of the issues and concerns commonly experienced by staff in institutions, and a summary of the topics covered in the research literature in the past, a review of documented studies pertinent to the research questions about tertiary teachers' perceptions of student evaluations/appraisals was presented. Perceptions held about teaching and learning and about evaluations/appraisals do seem to have an influence on how teachers respond to centralised evaluation systems, how they regard the data they collect and how they respond. Staff perceptions are bound up with institutional views and goals and the nature of the context within which teaching occurs. Such factors include whether the institution values, or is seen to value, teaching; whether there is an expectation that teachers' performance will be judged based upon results of student evaluations/appraisals; and the nature of the support for teachers when they receive evaluation/appraisal results. Practices around evaluations, including responding to and engaging with the feedback gained through evaluation/appraisals systems and processes, provide an indication of the perceptions teachers hold. 'Engagement' in this sense, is an active process. It involves multi-way communication about the feedback with a range of stakeholders. It incorporates reflection and action, professional development of teachers and the enhancement of teaching and learning. Thus 'engagement' with evaluation can entail both quality enhancement and teacher development.

To provide an overview of the New Zealand tertiary education scene, an environmental scan was presented of universities' and polytechnics' documents about evaluation available via the web. This scan, together with a brief summary of government expectations of universities and polytechnics, as they appear in the *Tertiary Education Strategy 2010-2015*, provided some baseline material to describe the context in which the current study occurred. Further detail was provided through a more in-depth description of the three participating institutions; the University of Otago, The University of Waikato (WU) and the Otago Polytechnic (OP).

Figure 2:1 presents a conceptual framework that pinpoints key components that have a bearing upon perceptions held by tertiary teachers about appraisals/evaluations and the nature

and extent of engagement with appraisals/evaluations. It illustrates the links between and amongst the theoretical ideas underpinning the study, drawn from the baseline material assembled in the literature review and environmental scan, and the shape of the report.

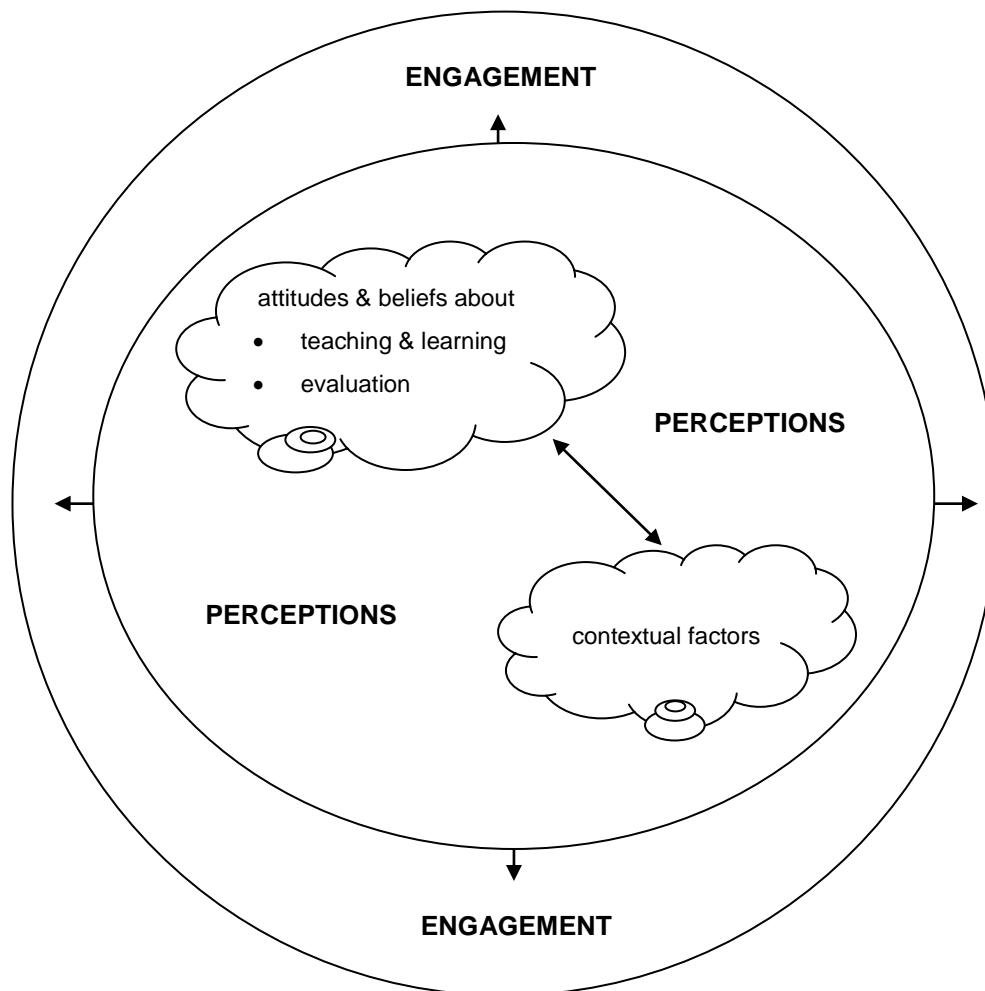


Figure 2:1 Influences on tertiary teacher engagement with student evaluations/ appraisals

Figure 2:1 shows that perceptions emerge from a combination of influences including views and beliefs about teaching and learning and the role of evaluation within the teaching-learning process. Contextual factors are significant; the role one plays in an institution and the institution’s expectations and demands for demonstration of effectiveness cause behaviours and views about oneself and one’s role. The perceptions one holds, therefore, play a large part in determining engagement with evaluation – practices, behaviours and motivations around evaluation. As Edström (2008) argues, “it doesn’t matter much what the

institution's intended purpose is. What is important is what the individual teachers *perceive* to be the purpose" (p. 100, emphasis in original).

In acknowledging that there can be a disparity between the public claims an institution makes about itself and the lived experience of individuals and groups making up that organisation, this study aimed to consider the mismatches between institutions' claims about student evaluations/appraisals and teachers' perceptions and lived experience of student evaluations/appraisals. Thus the overall research question became: *How do the current evaluation processes and practices influence teachers' thinking and behaviours in relation to student learning at all stages of the teaching and learning cycle?* The three research questions that formed the structure of the data gathering phases and also the analysis and reporting of the outcomes in this report match the elements shown in Figure 2:1. The three research questions were:

1. *What perceptions do tertiary teachers hold about student evaluations?* (refer to the central part of Figure 2:1)
2. *What factors (causes, influences) affect these views?* (again, refer to the central part of Figure 2:1)
3. *How do tertiary teachers engage with evaluation results and student feedback?* (refer to the outer circle in Figure 2:1)

Details of the research design are now presented in Chapter 3.

Chapter 3 : Research Design

3.1 Introduction

This chapter provides an overview of the research design. The research questions and objectives are revisited. Then the research approach is explained, followed by a description of the participants involved and the sources through which data were gathered. The analysis techniques used are then outlined. In this way, the basis is provided for the presentation of the results in the following chapters.

3.2 Aim, Research Questions and Objectives of the Study

This study aimed to consolidate, evaluate and add to past studies of student evaluations of teaching by investigating, from a New Zealand perspective, the extent to which teacher thinking and practices are currently informed by learner feedback provided in the formal institutional evaluation system.

Correspondingly, the broad research question was: *How do the current evaluation processes and practices influence teachers' thinking and behaviours in relation to student learning at all stages of the teaching and learning cycle?*

A number of more specific research questions were developed in the light of the background material presented in Chapter 2. These questions provided the conceptual framework for the study. The specific research questions assisted in planning the project in practical terms and allowed clear implementation targets to be set. The more specific research questions were:

1. *What perceptions do tertiary teachers hold about student evaluations?*
2. *What factors (causes, influences) affect these views?*
3. *How do tertiary teachers engage with evaluation results and student feedback?*

Answers to these questions would contribute to the achievement of findings in response to the broad research question.

As a way of setting clear targets for the project, to achieve the overall aim and to explore the research questions, the following became the objectives and thus gave shape to the project plan in practical terms:

1. To explore tertiary teachers' perceptions about student evaluations.
2. To identify how those perceptions impact on teacher thinking and practices.
3. To explore how tertiary teachers make use of information obtained from student evaluations at all stages of the teaching and learning cycle.
4. To compare teacher thinking and behaviours around evaluations across three participating institutions.
5. To make recommendations about evaluation processes so that evaluations can be used optimally for teaching development and enhancing student learning.

3.3 Research Plan

The research plan, organised into phases and linked directly to the objectives in section 3.2, is now described. Within each phase the general research activity that took place is explained (see also Appendix 4 for the timeline of project activity). More specific information about the data sources and their analysis appears in sections 3.4.2 and 3.4.3.

Phase 1 – development of a questionnaire and Phase 2 – administration of the questionnaire (linked to objectives 1, 2 and 3)

The literature search and environmental scan were undertaken as part of the first phase of this project (see section 2.1). The literature search summarised relevant international and national research on the uses of student evaluation of teaching and courses, and the environmental scan of institutional statements and policies of the eight universities and the 20 polytechnic institutions drew together examples of how student evaluations are used in New Zealand tertiary institutions (see section 2.2). This review process helped to establish the current thinking concerning the uses of student evaluation in tertiary contexts and provided the basis from which the questionnaire was devised. Upon reflection of this broader scene, more detailed information about the three participant institutions was documented, with a particular focus upon the evaluation/appraisal system, expectation and demands (see section 2.3). The researchers' knowledge of their institutional contexts thus also had a bearing on the content and presentation of the questionnaire, which appears in Appendix 1, the three participating

institutions and aimed to elicit tertiary teachers' perceptions on how student appraisals/evaluations influence teaching and learning processes, and their impact on both teacher development and student learning.

Details about the questionnaire, including its content, structure and administration, appear in section 3.4.2.1 and information about the analysis of the questionnaire data appears in section 3.4.3.1.

Phases 1 and 2 of the study occurred during February 2010 and May 2010.

Phase 3 – questionnaire data analysis, focussing on individual institutions (linked to objectives 1, 2 and 3)

In this phase, analysis of both the qualitative and quantitative questionnaire data was undertaken by each participating institution. The focus of this stage was on the analysis of individual institutional responses. The results of the analysis of the quantitative and qualitative questionnaire data can be found in Appendix 5 to Appendix 13. Discussion and presentation of all data in the light of the study's aim and research questions appear in Chapter 4 and Chapter 5.

Phase 4 – questionnaire data analysis, combined institutions (linked to objective 4)

Following analysis per institution, qualitative and quantitative questionnaire results were combined to assess the extent, pattern and nature of the impact of student appraisal/evaluation across institutions, and to identify any differences or similarities in results among the institutions. Relationships between themes emerging from qualitative and quantitative data were also examined in this light.

Descriptive statistics were generated, including frequencies, mean scores and standard deviations to identify trends and patterns across the three institutions. Since the sample was drawn from similar populations (tertiary teaching staff), the analysis examined relationships and significant differences between the mean scores of the three tertiary institutions. Further detail about the analysis of the questionnaire data can be found in section 3.4.3.1.

The results of the analysis of the combined institutional quantitative and qualitative data can be found in the Appendices. Discussion and presentation of all data, in the light of the study's aim and research questions, appear in Chapter 4 and Chapter 5.

Phase 5 - interviews (linked to objectives 1, 2 and 3)

In this phase, interviews were held as a way of finding out on a one-to-one basis about the responses provided to the questionnaire prompts. The interviews also investigated some aspects that had not been directly sought in the questionnaires, but which had emerged from the questionnaire responses. Further information about the interviews can be found in section 3.4.2.2.

The analysis of the interview data was similar to the analysis of the open-ended questionnaire data, and resulted in the identification of key themes. Information about the analysis can be found in section 3.4.3.2.

Discussion and presentation of all data, including the interview data, in the light of the research questions, appear in Chapter 4 and Chapter 5.

Phase 6 – combined institutional analysis, questionnaire and interview data (linked to objective 4)

In this phase, the findings from the three institutions were combined, considered, and then examined in the light of the literature review and environmental scan, focussing on tensions between the institutional use or expectations of appraisals/evaluations and their use by teachers for professional development. This drawing together of all findings served to highlight the ways tertiary teachers in the three institutions say they do or do not use evaluations/appraisals to improve their teaching practice and enhance student learning. The analysis provided the basis for further consideration of conclusions, implications and recommendations for the participating institutions and for all New Zealand tertiary institutions (phase 7 of the study).

Detail about this stage of the data analysis appears in section 3.4.3.1. The discussion of the combined analysis of all data, for each participating institution and overall, in the light of the research questions, appears in Chapter 4 and Chapter 5.

Phase 7 – conclusions implications and recommendations (linked to objective 5)

During this final phase of the project, data from all the previous phases were reviewed to examine the relationship between perceptions of evaluation/appraisal as found in this study, and institutional claims about the place and worth of evaluations, as found earlier in the

descriptions of the participating institutions appearing in section 2.3. Connections that emerged from the data were also reviewed in the light of the literature review, in section 2.1. This provided the essential evidence base for the recommendations for change, development and further research that institutions and teachers may initiate to use evaluations/appraisals to support and promote teaching and learning in all New Zealand tertiary institutions. The discussion of all outcomes of the study, in the form of a series of assertions, appears in Chapter 5 and recommendations appear in Chapter 6.

3.4 Macro Research Design - Structure and Procedures

The nature of the research question indicated that it was important to investigate the meanings that tertiary teachers attached to student evaluation/appraisal systems, policies and processes, which resulted in their various responses (perceptions and actions). Thus, the overarching research approach utilised for this study was interpretivist (Erickson, 1998). This interpretivist approach enabled us to highlight “the meanings and purposes attached by [tertiary teacher] human actors to their [evaluation related] activities” (Guba & Lincoln, 1989, p. 106).

An assumption of interpretivist methodologies is that there is an interactive link between researcher and research participant and that the values of the investigator influence the investigation (Guba & Lincoln, 1994). The interactive aspect suggests communication between the researcher and participant, and this communication, necessarily, is dialogic. The resultant exposition of perspectives, meanings, views and understandings leads to “more informed consciousness” (p. 110) about the phenomenon or situation under investigation, by both researcher and participant, thus supporting critical reflection and setting the scene for change and future action.

A questionnaire was used to gain broad insights into a large group of tertiary teachers’ perspectives. Those insights were then investigated in more depth, and on a personal level, through one-to-one interviews. In line with interpretivist approaches, as the project proceeded and the data became evidence (Miller & Fredericks, 2003), the research team held regular discussions about its developing understandings of the data in the light of the literature, the ongoing analysis and the institutional contexts under investigation.

Researchers, together with the other participants within a research context, interpret events and meanings. A consensual understanding is the aim (Guba & Lincoln, 1994). While an

account of events as they are lived by the participants is the aim of any research report of this type, it is the researchers who prepare such a report. It is ultimately the researchers' exposition of the other participants' interpretations that forms the final story of their meaning making and their experiences (Schwandt, 1994, 2007).

In interpretivist studies, the experienced lives of the actors within a context are described from their points of view. It is therefore important to describe the researchers' perspectives or frames of reference that background this study, to declare our particular interests and foci, and to provide a rationale for the way the participants' accounts and interpretation of their experiences were elucidated during the process of the study. A description of both the participants and the researchers follows.

3.4.1 Participants and researchers

In this section, the participants in the study are described. In line with the critical interpretivist approach, we include a section on the researchers, as well as those whom we invited to provide us with viewpoints and ideas through the data-gathering activity.

3.4.1.1 The tertiary teachers

The participants in the study were tertiary teachers from the UO, WU and the OP.

To ensure the anonymity of respondents, it was agreed between the three institutions that when identifying participants, a census approach would be most useful. This would allow for the collection of a wide range of views and perceptions, which would match the needs of the research project. Identification of teaching staff was not straightforward, as access to institutions' databases, including staff designations, required liaison with human resources sections. The process differed across the three institutions, each determined by institutional protocols. As an example, Appendix 2 shows an excerpt from project research notes documenting the process that was undertaken to identify prospective respondents at the University of Otago. Table 3:1 shows the number of staff who were invited to participate in the study.

Table 3:1 Number of Tertiary Teachers Invited to Participate in the Study from across the Three Institutions

Institution	Total Staff Surveyed
--------------------	-----------------------------

University of Otago (OU)	1443
The University of Waikato (WU)	663
Otago Polytechnic (OP)	320
All three institutions (ALL)	2426

While all teaching staff from the three participating institutions were invited to contribute, those who did were self-selected volunteers, with a willingness to share their views, positive or negative. Presumably, they had engaged in at least some thinking about evaluation and about evaluation of their teaching.

In line with the interpretivist approach being taken, the project was formulated in such a way as to investigate perceptions through a number of data sources, two of which involved gathering data directly from the tertiary teacher participants.

3.4.1.2 The researchers

Because we are products of our personal and social histories and of the contexts in which we work and live, we acknowledge that we cannot conduct research in a value-less way (Schwandt, 2007). All aspects of this study have been influenced by our beliefs, practices and experiences and each is the result of interpretation. Yet we, as researchers, maintain that our interpretations are not simply subjective viewpoints. It is therefore important to describe our perspectives or frames of reference that background this study, to declare our particular interests and foci, and to provide a rationale for the way the participants' accounts and interpretation of their experiences were elucidated during the process of the study.

The research team conducted this study against a backdrop of shared beliefs and practices about teaching and learning and about evaluation, which were borne out of personal and group experiences. The views of the researchers were also based on knowledge of the wider research literature and evidence gathered through the data sources and the processes of developing understanding about the data we were collecting. At the time of the study, the research team was made up of:

- two academic staff developers – Dorothy and Sarah – both academic staff members of universities, with a number of years' experience in education and teacher professional development in New Zealand and either South Africa or Australia

- one research assistant – Lynley – with a number of years’ experience working in universities on a variety of research projects on a range of topics, but most notably in the areas of marketing education research and broader tertiary education research
- three evaluation-specific roles – Jo, Trudy, Stuart – each with his or her own professional background (law, engineering, commerce.), and with a number of years’ experience and responsibility for the oversight and administration of the student evaluation/appraisal of teaching and course processes in universities or polytechnics.

As staff of our respective institutions whose main roles include supporting other staff in the broad areas of teaching, learning, evaluation and teacher professional development, we were particularly interested in investigating the various responses of our colleagues when faced with our institutions’ centralised student evaluation/appraisal systems, processes and results. As stated earlier, in section 2.1.1, in our work with staff we knew of the reality of the many responses, experiences, emotions, views and beliefs and had heard many myths and anecdotes about evaluation. We had experienced many similar responses ourselves in our own teaching and administrative work. Thus, because our interests lay predominantly with assisting our colleagues and our institutions in their translation of theoretical notions and institutional policies about evaluation into meaningful entities for themselves, we were keen to investigate the tensions that we were aware of between the developmental purposes for evaluation and the auditing or quality assurance purposes. We were interested in bringing to light the perceptions teachers hold about student evaluation and the impact evaluation systems and regimes in tertiary institutions have on those perceptions and associated behaviours. We were also interested in the detection of underlying themes, structures and platforms that teachers may have found useful in helping them make sense of their evaluation experiences.

3.4.2 Data sources

In this section, each of the data sources is described. Details about the processes involved in administration and data collection are presented. The main data sources were a questionnaire and interviews.

3.4.2.1 Questionnaire

Phase 1 of the study included the development of a questionnaire. As stated in section 3.3, the intention of the questionnaire was to elicit tertiary teachers' perceptions on how student evaluations/appraisals influence teaching and learning processes, and their impact on both teacher development and student learning. As well as questions that prompted respondents to report their perceptions, demographic information was collected. A question (Q33) asking for volunteers to be interviewed at a later stage was also included (phase 5 of the study). The questionnaire used both Likert scale questions and open response questions, to allow both quantitative and qualitative analysis.

A small pilot was run at each of the institutions. The example in Figure 3:1, taken from project research notes, serves to illustrate how this pilot occurred at the University of Otago.

Pilot Survey (Research Notes University of Otago, 2010)

- At the end of March we sent out an invitation to complete a pilot survey asking for feedback about the questions, format etc. The pilot was open from Tue 30 Mar to Thu 8 April, 2010.
 - It was sent to 45 staff of which 34 were HEDC staff (incl PG students). The rest were external to HEDC and included staff from Distance Ed, Medical Education, Anthropology, Gender and Sociology, Maths & Stats, Marketing, Physics, Info Science.
 - 23x responded and a number of changes were made from their feedback. After liaising with Waikato and OP more changes were made until we ended up with the final version.
-

Figure 3:1 Research notes about the questionnaire pilot – an example

The resulting questionnaire, which appears in Appendix 1, was modified for each of the institutions participating in the study to ensure that terminology and contextual references were appropriate for each setting. It consisted of four broad sections: Section A explored current practices; Section B explored perceptions of the data and influence on practice; Section C asked for demographic information; and the final Section D asked about interview availability.

During the questionnaire development, it became apparent that answering some questions may have been more difficult for teachers with little experience to draw upon, particularly the questions in Section A about current use. However, considering the focus of the study, perceptions from all teaching staff, no matter their experience, was deemed to be important.

For this reason, and to acknowledge that inexperience does not necessarily mean lack of knowledge, respondents who answered ‘no’ to both parts of Q1 (*Have you ever run student evaluations/appraisals using the centralised system of evaluation/appraisal at: (a) Another tertiary institution? and (b) Your current institution?*) were asked to complete the demographic section of the questionnaire and to make general comments if possible. As a result, many who answered ‘no’ to both parts of Q1 did continue and provided responses to some of the questions, including the demographics section.

Once the questionnaire development was complete, phase 2 of the project began. Academic teaching staff at the Universities of Otago and Waikato and the Otago Polytechnic were invited to complete the questionnaire, which was administered online. An invitation to complete the online questionnaire was sent to all identified teaching staff. Identification of teaching staff was not straightforward, as access to institutions’ databases, including staff designations, required liaison with human resources sections. The process differed across the three institutions, determined by institutional protocols. As an example, Appendix 2 shows an excerpt from project research notes documenting the process that was undertaken to identify prospective respondents at the University of Otago. In addition, as noted in Appendix 2, to reduce bias potential participants were randomly split into three groups, A, B and C. This was done by adding a random number to each staff member in the spreadsheet and then sorting on that number to reorder them all. Each group was sent a questionnaire in which the order of the options for Q2, Q4 and Q7 was changed.

Respondents were given three weeks to complete and submit their responses. At each institution, the questionnaire went live on or around Thursday 22 April 2010. The first reminder was sent on or around Monday 3 May 2010. The last chance reminder was sent on or around Tuesday 11 May 2010. The questionnaire was closed on or around Thursday 13 May 2010. The dates varied slightly for each institution, but only by a day or two. There were some returned emails that were not valid and staff advising that they no longer worked at the institution (for example, recently retired). These were removed from the lists so the number of staff surveyed was slightly lower than the initial figure. The total staff surveyed figure below is this adjusted number, as it represents the staff who could be expected to complete the questionnaire.

Table 3:2 shows for each institution and overall, the number of teachers invited, the total responses received and response rates.

Table 3:2 Overall Distribution Report – Total Responses to Questionnaire

Date questionnaire run: late April/early May 2010 Duration: three weeks Format: online			
Institution	Total staff surveyed	Total responses	Response Rate
OU	1443	670	46%
WU	663	242	37%
OP	320	153	48%
ALL	2426	1065	

Demographic details of the respondents were presented in section 3.4.1.1.

Once the questionnaire had closed, collation of the data was undertaken at each of the three institutional sites. At this initial stage of analysis, quantitative data were entered into Excel and prepared for statistical analysis using SPSS. The summary of the quantitative questionnaire data appears in Appendix 5. Note that due to a technical problem, OP attracted a high number of nil responses for its randomly split group C (41 respondents). This affected questions 4(e) and 4(f) and all responses for OP group C were recorded as nil. (See note in Appendix 5). The qualitative responses from open-ended questions were also collated and entered into Excel, ready for later analysis. See further details about the analysis of the questionnaire data in section 3.4.3.1.

3.4.2.2 Interviews

Phase 5 of the study was the interview phase. In this phase, 60 interviews, 20 at each institution, were held as a way of probing the more general questionnaire responses. Volunteers recorded their interest in an interview via the questionnaire (see the final questionnaire prompts in Appendix 1). Table 3:3 shows the number of interview volunteers from each institution.

Table 3:3 Volunteers for Interview from each Participating Institution

Institution	Number of interview volunteers
OU	292
WU	121
OP	97

Twenty volunteers from each institution were selected for interview. Selection, using this convenience quota criterion sample, was purposive (Patton, 1990). Interviewees were selected to ensure there was a spread of:

- length of tertiary teaching experience
- range of disciplines/faculties/schools/divisions
- range of academic position – lecturer, senior lecturer, associate professor, professor
- promoted versus not promoted
- gender.

At each institution, the semi-structured interviews lasting about 30–60 minutes were conducted, audio recorded and later transcribed. The core interview questions were based on key themes identified in the questionnaire responses and asked interviewees to elaborate on their experiences and views of student evaluations/appraisals. The interview explored, in depth, teachers' thinking and attitudes and the behaviours related to student evaluations/appraisals in which they claimed they engaged. The core interview questions appear in Appendix 3.

To enhance quality assurance, the transcripts of the interviews were returned to the interviewees for checking and comment. Interviewees were given this option, although many did not wish to see transcripts.

See section 3.4.3.2 for information about the analysis of the interview data.

3.4.3 Analysis techniques

3.4.3.1 Questionnaire

During phase 3 of the study, the first major analysis phase, the focus was on the questionnaire responses gathered at each institution. At each institutional site, the quantitative data were used to look for correlations between the scores of the individual question responses to show to what extent the process of doing a student evaluation/appraisal impinges on a teacher's choice of teaching strategies, including if, and how, evaluation may constrain innovation in teaching. Using the demographic data, the correlation scores of these main responses were examined to look for factors influencing them, such as years of teaching in the tertiary sector and levels of employment. Demographic data on age, ethnicity and gender were collected, to demonstrate the range of the participants.

3.4.3.1.1 Quantitative data

Because of the variation in nomenclature across the three institutions, for practical analysis purposes, responses from Q28 *What position do you currently hold?*, were grouped in a way that best reflected the principal activity/ies undertaken as part of the positions, as related to teaching and evaluation. The positions were grouped into the following categories (see Table 4:7 in section 4.2).

- 1) *Junior teaching positions*, which included Professional Practice Fellows, Teaching Fellows, Senior Teaching Fellows, Tutors and Senior Tutors (approximately 14 per cent).
- 2) *Lecturer teaching positions*, which included Lecturers and Senior Lecturers (approximately 56 per cent).
- 3) *Senior teaching positions*, which included Principal Lecturers, Associate Professors, Professors, Programme Managers/Coordinators, Heads of Schools, Senior Managers (approximately 27 per cent)
- 4) *Other*, which included the remaining three per cent (30 responses), the designations of which were not able to identified.

It is acknowledged that these groupings are fairly crude insofar as they do not necessarily reflect the reality that actual activities of role holders can vary tremendously and can change over time, even within the same designated role. For example, an individual Senior Teaching Fellow, grouped in the 'junior teaching position' category, may be involved in research as well as teaching, and teaching for that person may involve extensive or minimal administration and management responsibilities. While in the universities the scope of the 'lecturer teaching position' and the 'senior teaching position' roles usually do involve teaching, research and service, at the polytechnics the breadth and extent of research and teaching is not necessarily comparable with the universities.

A One-way Analysis of Variance (ANOVA) was used to investigate differences in the means scores between the three intuitions on all the Likert scale questions. When looking at differences between institutions, using the combined data, it is possible to test differences between mean values using one-way analysis of variance (ANOVA). When assessing the mean differences for demographics and the Likert scale-type questions, the analysis of variance was undertaken on each institution's individual data. The one-way ANOVA for independent samples assumes that the scale on which the dependent variable is measured has the properties of an equal interval scale and the samples have approximately equal variances. Other assumptions relate to normal distributions, equal variances and independence.

The Kruskal–Wallis test is a non-parametric alternative to a one-way ANOVA and is based on the ranks of the data. It tests that the three medians are equal or at least one median is different from others. The Kruskal–Wallis is used alongside the ANOVA to test whether the mean and medians are different between groups. The Kruskal–Wallis in this report is used to support the claim based on ANOVA where there are significant differences in means between institutions and between groups.

Cross tabulations were used to test the association between a row and a column in a two-way table. In using a Chi-square analysis, researchers are interested in the frequency with which individuals fall in a category or combination of categories such as, in this study, whether they had been promoted, the institutions they were from and years of teaching in a tertiary institution. For example, for Q2, the Pearson Chi Square was used to measure the divergence of observed data from expected values, and to test if there was significant proportion of the participants saying 'yes' as opposed to saying 'no'. So for Q2 *Please identify why you use*

student evaluations/ appraisals with a ‘yes’/‘no’ statement, the proportion of males to females, proportion of staff who had been promoted or not promoted and if there was a difference between observed and expected values, were able to be determined.

3.4.3.1.2 Qualitative data

The analysis of the qualitative data from the open-ended section of the questionnaire involved searching for themes, using a constant comparative technique (Dye *et al.*, 2000; Silverman, 2001). The key ideas and issues identified through the literature review contributed to this analysis. In particular, the process involved seeking descriptive and theoretical ties between teachers’ perceptions of the place and worth of evaluations: their beliefs about their institution’s view of evaluations, as well as their reported experiences of how they used them in, and for, their teaching. For each open-ended question in Section B of the questionnaire, a list of key themes and sub-themes, supported by relevant quotations taken from the responses, was the result. Codes were generated, linked to key themes, and entered into an Excel spreadsheet. In addition, a tally was made of the number of responses made to the open-ended questions, across institutions. This tally can be seen in Table 3:4, below.

Table 3:4: Summary of Response Rates – Comment Questions

	Total responses	Q10	Q12	Q14	Q16	Q18	Q20	Q21	Q22	Av Tot
OU	670 (100%)	472 (70%)	432 (64%)	421 (63%)	388 (58%)	421 (63%)	396 (59%)	455 (68%)	276 (42%)	408 (61%)
WU	242 (100%)	199 (82%)	183 (76%)	182 (75%)	174 (72%)	185 (76%)	179 (74%)	204 (84%)	102 (42%)	176 (73%)
OP	153 (100%)	46 (30%)	39 (25%)	42 (27%)	40 (26%)	68 (44%)	59 (39%)	120 (78%)	65 (42%)	60 (39%)
All	1065 (100%)	717 (67%)	654 (61%)	645 (61%)	602 (57%)	674 (63%)	634 (60%)	779 (73%)	443 (42%)	644 (60%)

The results of the analysis appear within the discussion in Chapter 4 and Chapter 5. The full set of comment data analysis for these questions, their theme and sub-theme coding, examples of comments categorised within themes and sub-themes and theme coding by institution can be found in Appendix 6 to Appendix 13.

3.4.3.2 Interviews

The analysis of the interview data took place during phase 5 of the project, first on an individual institution basis, then combined (phase 6). The analysis process was similar to the analysis of the open-ended questionnaire data (see phase 3 and section 3.4.3.1.2), and resulted in the identification of key themes.

The coding, which resulted in a set of core themes, engaged a constant comparative technique (Dye *et al.*, 2000; Silverman, 2001). Themes that emerged as evidence about the major trends in the interview data were compared and contrasted with questionnaire data and literature review topics. The result was the development of a framework that was then used both for a closer analysis of interview data by institution and for the analysis of the combined institutional interview data. The core themes making up this framework became:

- teaching and learning beliefs (with particular focus on responses to Q2a of the interview questions: *How do you see yourself as an educator?*)
- students' capacity to make judgements (with particular focus on responses to Q2b of the interview questions: *Generally, do you think students are able to make judgements about the quality of teaching and their learning experience?*)
- personal/emotional factors (with a particular focus on responses to Q2c of the interview questions: *When you receive the results of evaluations from students, how do you feel?*)
- other factors (for example, timing, promotion, use *etc.*)
- engagement with evaluations.

Where relevant in the report of the findings in Chapter 4, a mix of quotations from the interviews and researcher notes are used. Quotations appear in italics and relevant researcher notes appear in square brackets alongside relevant quotations. Unlike quotations drawn from the questionnaire comment data, interview quotations are not accompanied by demographic details. The sections reporting interview findings are organised by institution, so that the institutional origin of the interviewee providing each quotation is clear.

3.5 Quality Assurance

Quality was assured and enhanced using a number of strategies and devices, similar to those recommended by Guba and Lincoln (1989). The overall approach incorporated regular team discussions and the ongoing development of assertions about the evidence gathered through the data sources. The development of the assertions occurred by integrating those regular discussions with checks (comparisons and contrasts) with the literature. In this way, assertions were proposed, refined, confirmed or refuted, then further refined in the light of the integrated processes of reflection, discussion, data and literature checks.

Where the interviews were concerned, interviewees were invited to review transcripts to confirm that their words and thoughts had been captured accurately in the ‘verbatim’ interview transcription. Of all the interviewees who replied to this invitation, one took the opportunity to add further notes. All others agreed that the transcriptions were accurate. Finally, in the report of the interviewees’ words, only general descriptive labels were used (for example, institution, position) and care was taken to remove any references that could reveal the identity of the interviewees.

During the analysis of the quantitative questionnaire data, the valid percentage was used from the frequencies tests, summing only the number answering the questions and omitting the missing data. The Kruskal–Wallis independent median test was undertaken to determine whether there was a median difference supporting the difference of mean tests so as to not violate any assumptions associated with the one-way analysis of variance and the use of categorical data.

3.6 Summary of Chapter 3

This chapter provided an overview of the research design, which included a description of the research approach followed by the design and methods. The participants and the researchers were described, as were the data sources and analysis techniques. Finally, quality assurance measures were described.

Chapter 4 : Findings

4.1 Introduction

In this chapter, the analysis of each of the data sources is drawn upon to provide evidence for claims made about tertiary teachers' views on student evaluations and the impact evaluation and evaluation systems, institutional contexts, and processes and procedures, may be having on practice. The chapter begins by looking at the demographics of the respondents who participated in the questionnaire, broken down by institution. The rest of the chapter is organised around the three questions that emanated from the literature review and which were refined and then illustrated in diagrammatic form as a conceptual framework in Figure 2:1.

To match the overarching qualitative nature of the study, the findings are presented in an integrated way around the three questions that emerged as the study progressed. The data from individual data sources are thus not presented as isolated entities in this chapter. By focussing on the research questions a more meaningful discussion is provided that better matches the intentions of the research and the study's aims and objectives, and which enables cross-comparison across the data sets.

Further discussion of the findings appears in Chapter 5, based on a series of assertions and including explicit links to the literature review. Major outcomes and a list of recommendations are then presented in Chapter 6.

4.2 The Demographics

As mentioned in Chapter 3, the questionnaire was run online during late April and early May 2010. The questionnaires were open for a period of three weeks, with reminders at the end of each week. Table 4:1 below, shows for each institution and overall the number of teachers surveyed, the number of responses received and the corresponding response rates.

Table 4:1: Overall Distribution Report – Total Responses to Questionnaire

Institution	Total staff surveyed	Total responses	Response rate
--------------------	-----------------------------	------------------------	----------------------

OU	1443	670	46%
WU	663	242	37%
OP	320	153	48%
ALL	2426	1065	44%

The demographics of the participants are presented in the following tables. The data were provided by the participants in answer to a series of demographic questions in the questionnaire. The question number from which the data were drawn is given in parentheses at the end of each table title. The questionnaire appears in Appendix 1. In the following tables, *n* is the number of participants who responded to the particular question. The data are presented firstly showing the number of respondents for a particular item. The response rate for all demographic questions ranges between 82 per cent and 99 per cent.

Table 4:2: Summary of Participant Gender/Sex (Q23)

Institution	<i>n</i>	male	female
OU	584	314 (54%)	270 (46%)
WU	236	112 (47%)	124 (53%)
OP	137	43 (31%)	94 (69%)
Total	957	469 (49%)	488 (51%)

n is the number of participants who responded to the particular question

Table 4:3: Summary of Participants' Ethnicity Data (Q24)

Inst	<i>n</i>	NZ/ Pakeha	Māori/ NZ of Māori descent	Pasifika	Asian	Non-NZ European/ Caucasian	other
OU	565	367 (65%)	12 (2%)	4 (1%)	31 (5%)	122 (22%)	29 (5%)
WU	239	141 (59%)	16 (7%)	1 (0%)	11 (5%)	52 (22%)	18 (8%)
OP	125	108 (86%)	4 (3%)	0	0	13 (10%)	0
Total	929	616 (66%)	32 (3%)	5 (1%)	42 (5%)	187 (20%)	47 (5%)

n is the number of participants who responded to the particular question

Table 4:4: Summary of Participants' Age Data (Q25)

	<i>n</i>	<= 30	31-35	36-40	41-45	46-50	51-55	56-60	61 +
OU	588	19 (3%)	43 (7%)	64 (11%)	79 (13%)	124 (21%)	113 (19%)	72 (12%)	74 (13%)
WU	239	5 (2%)	17 (7%)	26 (11%)	31 (13%)	46 (19%)	39 (16%)	37 (15%)	38 (16%)
OP	136	3 (2%)	13 (10%)	20 (15%)	22 (16%)	23 (17%)	27 (20%)	12 (9%)	16 (12%)
Total	963	27 (3%)	73 (8%)	110 (11%)	132 (14%)	193 (20%)	179 (19%)	121 (13%)	128 (13%)

n is the number of participants who responded to the particular question

Table 4:5: Summary of Participants' Tertiary Teaching Experience (Q26)

Yrs Inst	<i>n</i>	0-5	6-10	11-15	16-20	21+
OU	590	108 (18%)	119 (20%)	128 (22%)	95 (16%)	140 (24%)
WU	239	32 (13%)	49 (21%)	56 (23%)	47 (20%)	55 (23%)
OP	137	40 (29%)	38 (28%)	23 (17%)	15 (11%)	21 (15%)
Total	966	180 (19%)	206 (21%)	207 (21%)	157 (16%)	216 (22%)

n is the number of participants who responded to the particular question

Table 4:6: Summary of the Nature of Employment of Participants (Q27)

Institution	<i>n</i>	Confirmation path	Permanent/continuing	Fixed term	Casual/other
OU	586	137 (23%)	396 (68%)	42 (7%)	11 (2%)
WU	240	n/a	213 (89%)	26 (11%)	1 (0%)
OP	134	n/a	114 (85%)	12 (9%)	8 (6%)
Total	960	137 (14%)	723 (75%)	80 (8%)	20 (2%)

n is the number of participants who responded to the particular question

Table 4:7: Summary of Positions Held by Participants in their Institutions (Q28)

Institution	<i>n</i>	Junior teaching positions	Lecturer teaching positions	Senior teaching positions	Other
OU	582	95 (16%)	313 (54%)	161 (28%)	13 (2%)

WU	231	36 (16%)	120 (52%)	58 (25%)	17 (7%)
OP	134	n/a	99 (74%)	35 (26%)	0
Total	947	131 (14%)	532 (56%)	254 (27%)	30 (3%)

n is the number of participants who responded to the particular question

The type/classification/naming of positions varied slightly for the different institutions, but for most of the analysis the positions were grouped into the four categories appearing in Table 4:7, above. The positions included in each category are:

- *Junior teaching positions*: Professional Practice Fellow, Teaching Fellow, Senior Teaching Fellow, Tutor and Senior Tutor
- *Lecturer teaching positions*: Lecturer and Senior Lecturer
- *Senior teaching positions*: Principal Lecturer, Associate Professor, Professor, Programme Manager/Coordinator, Head of School and Senior Manager.

Table 4:8: Percentage of Participants who had Applied/Not Applied for Promotion During the Previous Five Years (Q29)

Institution	<i>n</i>	yes	no
OU	579	264 (46%)	315 (54%)
WU	235	154 (66%)	81 (34%)
OP	138	61 (44%)	77 (56%)
Total	952	479 (50%)	473 (50%)

n is the number of participants who responded to the particular question

Table 4:9: Percentage of Participants who had been/had not been Promoted During the Previous Five Years (Q30)

Institution	<i>n</i>	yes	no
OU	582	268 (46%)	314 (54%)
WU	235	138 (59%)	97 (41%)
OP	139	68 (49%)	71 (51%)
Total	956	474 (50%)	482 (50%)

n is the number of participants who responded to the particular question

Table 4:10: Summary of Division/Faculty/School in which the Participants were Predominantly Based (Q31)

Institution	n=	Commerce	Health Sciences	Humanities	Science	Other
OU	582	53 (9%)	243 (42%)	160 (27%)	120 (21%)	6 (1%)
WU	235	51 (22%)	0	125 (53%)	55 (23%)	4 (2%)
OP	132	13 (10%)	33 (25%)	42 (32%)	36 (27%)	8 (6%)
Total	949	117 (12%)	276 (29%)	327 (34%)	211 (22%)	18 (2%)

n is the number of participants who responded to the particular question

As shown in Table 4:10, at WU the seven faculties/schools were split into four discipline areas similar to the UO. Therefore, participants from the Faculties of Arts and Social Sciences, Education, Law and the School of Māori and Pacific Development were placed under Humanities.

Table 4:11: Summary of Location where the Participants were Predominantly Based (Q32)

OU	<i>n</i>	Dunedin	Wellington/Auckland	Christchurch	Invercargill	Other
	584	489 (84%)	49 (8%)	44 (8%)	2 (0%)	0
WU	<i>n</i>	Hamilton	Tauranga			Other
	238	230 (97%)	4 (2%)			4 (2%)
OP	<i>n</i>	Dunedin	Distance (based at home)	Central Otago	Wintec	Other
	133	108 (81%)	10 (8%)	9 (7%)	4 (3%)	2 (2%)

n is the number of participants who responded to the particular question

4.3 The Questions

The literature review highlighted research and thinking around evaluation in relation to: planning courses and teaching approaches; receiving and responding to feedback; and the use of feedback for professional development. Three core contributing questions addressing the key issues found in the literature were developed as part of the process of refining the data-gathering tools for this study:

1. *What perceptions do tertiary teachers hold about student evaluations?*
2. *What factors affect these views?*
3. *How do tertiary teachers engage with evaluation?*

These contributing questions are now used to frame the presentation of the data and to provide an orientation to, and basis for, the discussion of the assertions in the next chapter.

4.3.1 Question 1: What perceptions do tertiary teachers hold about student evaluations?

In this section, the findings that provide insights into the perceptions about evaluations held by the tertiary teachers who participated in the study are presented. As highlighted in the literature review, and shown diagrammatically in Figure 2:1, the perceptions teachers hold about evaluations underpin and drive behaviours. These perceptions are determined by beliefs about teaching, learning and evaluation, prior experiences and the contexts in which teachers operate. To address the first contributing research question – *What perceptions do tertiary teachers hold about student evaluations?*– data were drawn together from the questionnaire prompts, Q2, Q4 and Q17, together with comment data from Q3, Q5 and Q18, as well as from the interviews.

4.3.1.1 Teachers’ perceptions of student evaluations – questionnaire responses

Table 4:12 shows that in answer to Q17 *Do you personally think it is worthwhile to gather student evaluation/appraisal data about teaching and courses/papers?*, of the number of responses to this question (n=943), just under three quarters of respondents from across all institutions claimed that they regard gathering evaluation/appraisal data as personally worthwhile (73 per cent responding with a one or a two). Of the institutions, OP responded with the highest frequency of ones and twos at 85 per cent, followed by WU with 74 per cent and finally OU with 70 per cent.

Table 4:12 Frequency of Responses to Q17 (Do you personally think it is worthwhile to gather student appraisal/evaluation data about teaching and courses/papers?)

	very worthwhile (1&2)	middle ground (3)	not at all worthwhile (4&5)
ALL (combined institutions) Mean = 1.92; Median = 2	73%	16%	11%
OU Mean = 2.03; Median = 2	70%	18%	12%
WU Mean = 1.86; Median = 1	74%	15%	11%

OP			1	
Mean = 1.55; Median = 1	85%		10%	5%

Indications of the reasons why evaluations/appraisals are positively regarded in general, can be gleaned from an examination of the responses to Q18, which asked for an explanation about Q17 responses (Q18 *Please explain your answer to Q17*). In the Q18 comments, the terms ‘evaluation’ and ‘appraisal’ are not always used by participants to refer to the centralised student evaluation system run in their institutions. Sometimes their comments refer to broader evaluation practices. The comments provide an indication of the scope of general perspectives that teachers hold about evaluation and illustrate why many regard evaluations/appraisals personally worthwhile.

Two themes accompanied by a series of sub-themes were identified in the responses to Q18, highlighting factors that either enhance or limit teachers’ sense of worth of student evaluations as contributing to educational processes (see Appendix 10 for descriptive detail about Q18 themes and sub-themes):

- *Q18 Theme 1* – factors enhancing teachers’ sense of worth of student evaluation data, and sub-themes
 - 1a. informs course/ teacher development
 - 1b. helps identify student learning needs/ experiences
 - 1c. informs and provides evidence for use in quality/ summative/performance-based processes
 - 1d. forms part of a range of evaluation practices
 - 1e. has some worth/importance.
- *Q18 Theme 2* – factors limiting teachers’ sense of worth of student evaluation data, and sub-themes
 - 2a. use for quality/summative/performance-based processes
 - 2b. other evaluation methods better/preferred
 - 2c. current system limitations
 - 2d. quality of student responses questionable
 - 2e. teachers judged on factors outside their control

- 2f. difficulties in relation to interpretation and how to use the data effectively
- 2g. generally not valued/useful.

It can be seen from Appendix 10 that amongst the factors contributing to (Q18 Theme 1) teachers' sense of educational worth of evaluations, sub-themes 1a (*evaluations are personally worthwhile to inform course/teacher development*) and 1b (*evaluations are worthwhile because they help identify student learning needs/experiences*) formed the highest percentage of comments (19 per cent each, 38 per cent added together). Where sub-theme 1a, *inform course/teacher development*, is concerned, comments included, for example, reference to the insights that an individual can develop by gaining others' viewpoints.

Keeps you on your toes - review by others is a great way to identify how others see you... and how they see your strengths and weaknesses. (Q18 sub-theme 1a, OU, lecturer teaching position, 16–20 years' tertiary teaching experience, permanent, Health Sciences.)

I have no other 'thermometer' on a regular basis to judge where my teaching is at! We do have observations made of our teaching by a senior/teaching coordinator but student evaluations are very valuable to me to give my teaching reflections balance. (Q18 sub-theme 1a, WU, other category for position, 6–10 years' tertiary teaching experience, continuing, Humanities.)

... on the whole they [evaluations] do provide valuable information to refine how I do things in class. (Q18 sub-theme 1a, OP, senior teaching position, 21+ years' tertiary teaching experience, permanent, Commerce.)

Other respondents focussed on the worth of evaluations for providing information about students' learning experiences (*Q18 sub-theme 1b help identify student learning needs/experiences*). Some examples of comments within this sub-theme include:

Teaching is a conversation and a good conversationalist understands the importance of listening. Students have a right to have their voice heard; and teaching is about helping students to find their voice. (Q18 sub-theme

1b, OU, lecturer teaching position, 11–15 years' tertiary teaching experience, permanent, Sciences.)

I know it can seem a rote exercise, but it is actually the one time a student can say what they think anonymously and mostly students take the option seriously and offer valuable feedback. (Q18 sub-theme 1b, WU, senior teaching position, 16–20 years' tertiary teaching experience, continuing, Humanities.)

[Evaluation] is a great way to learn about what is good and what is bad - hard to be totally objective about your methods etc. and students are great at honesty in this forum! (Q18 sub-theme 1b, OP, lecturer teaching position, 0–5 years' tertiary teaching experience, permanent, Sciences.)

Appendix 10 also shows that amongst the (Theme 2) factors limiting teachers' sense of educational worth of evaluations, the sub-themes 2a (*personal worth of evaluations is limited by their use for quality/summative performance-based processes*), 2c (*personal worth of evaluations is affected by limitations of the current evaluation systems*) and 2d (*personal worth of evaluations is limited by the questionable quality of student responses*) attracted the highest percentage of comments (seven per cent, 14 per cent and 10 per cent respectively).

Awareness that judgements and decisions are made on the basis of evaluation data without close knowledge of the teaching context about which the data were gathered was of concern to the two university respondents who made comments around sub-theme 2a, but much less so for OP respondents. For example:

I do not however, approve of the institution's tendency to use them [evaluations] as weapons against staff. A heavy-handed hierarchical approach from academics with little knowledge of the course or the students and even less interest in teaching or its context, is counter-productive. (Q18 sub-theme 2a, OU, lecturer teaching position, 16–20 years' tertiary teaching experience, permanent, Humanities.)

The rating questions are rather useless but perhaps useful for a promotion committee to make broad judgements. That is their sole value nothing else. The reason for that is that they do not specifically tell you what is wrong

or what is right. The comments do that best. Also as mentioned earlier the statistical rigour in many of these appraisals would make a real statistician seriously question their meaning. (Q18 sub-theme 2a, WU, lecturer teaching position, 11–15 years' tertiary teaching experience, continuing, Sciences.)

It is valuable, but only if taken in context! One negative comment and 100 positive ones is a very good result. However, management have a tendency to focus on that one comment. Often there are reasons other than the quality of the teaching for negative evaluations. (Q18 sub-theme 2a, OP, senior teaching position, 6–10 years' tertiary teaching experience, permanent, Health Sciences.)

Other comments around sub-theme 2c (*personal worth of evaluations is reduced by limitations of the current evaluation systems*) noted issues with central evaluation systems including the surveys themselves and the processes surrounding administration and analysis.

[Student evaluation is personally worthwhile, but] not when it is using Likert scale methods. I value qualitative data much more highly. (Q18 sub-theme 2c, OU, senior teaching position, 6–10 years tertiary teaching experience, permanent, Commerce.)

What is not good in our system is the standard format which is unsuited to so many diverse courses. (Q18 sub-theme 2c, WU, lecturer teaching position, 11–15 years' tertiary teaching experience, continuing, Humanities.)

The problem is, you tend to get evaluations from the students who attend class and it is probably not the target group. The ones who don't attend are the ones who might not enjoy the teaching methods. (Q18 sub-theme 2c, OP, lecturer teaching position, 6–10 years' tertiary teaching experience, permanent, Health Sciences.)

A particular concern of other respondents was with the responses provided by students. The following comments illustrate sub-theme 2d (*personal worth of evaluations is limited by the questionable quality of student responses*).

It gives you some feedback but students do not necessarily evaluate quality teaching and are influenced by general friendliness, easy marking standards, and 'flexibility' (i.e. lack of adherence) to deadlines, regulations, etc. - peer assessment would be more valuable. (Q18 sub-theme 2d, OU, lecturer teaching position, 16–20 years' tertiary teaching experience, permanent, Humanities.)

I have reservations about how students can be manipulated; how they enjoy good personalities and males more than females; that many don't like being challenged to think; that the value system against which they give feedback is often very different from my own (e.g. they might say something was boring); and that they do not understand that the feedback is treated as seriously as it is by the institution - or myself. There is no accountability for and by students in the student feedback system. (Q18 sub-theme 2d, WU, senior teaching position, 21+ years' tertiary teaching experience, continuing, Humanities.)

I often feel that the students put no thought into their evaluations. (Q18 sub-theme 2d, OP, lecturer teaching position, 21+ years' tertiary teaching experience, permanent, Commerce.)

More detail concerning teachers' perceptions about student evaluations can be seen through an examination of the responses to questions 2, 3, 4 and 5 of the questionnaire.

Question 2 *Please identify why you use student evaluations/appraisals?* was made up of the lead question followed by eight prompts, each of which requested a 'yes'/'no' answer. Table 4:13 presents the overall Q2 results across all institutions.

Table 4:13 Overall Responses to Q2 (Please identify why you use student evaluations/appraisals)

Q2: Why use evaluations/appraisals? 'yes' responses	OU	WU	OP	Total respondents
--	-----------	-----------	-----------	--------------------------

Total	561	227	136	924
<i>2a Because it is required</i>	n=462 82%	n=193 85%	n=132 97%	n=787 85%
Total	570	215	130	915
<i>2b For my professional development</i>	n=499 88%	n=179 83%	n=115 88%	n=793 87%
Total	565	225	108	898
<i>2c For my promotion</i>	n=459 81%	n=191 85%	n=70 65%	n=720 80%
Total	508	203	109	820
<i>2d For confirmation/PGS/salary review</i>	n=338 67%	n=130 64%	n=76 70%	n=544 66%
Total	569	220	136	925
<i>2e For course/paper refinements</i>	n=507 89%	n=191 87%	n=128 94%	n=862 89%
Total	591	227	141	959
<i>2f To get feedback on student learning experiences</i>	n=543 92%	n=216 95%	n=135 96%	n=894 93%
Total	461	202	104	767
<i>2g To provide feedback to my students</i>	n=190 41%	n=86 43%	n=65 63%	n=341 44%
Total	452	195	110	757
<i>2h To report on quality matters</i>	n=172 38%	n=92 47%	n=81 74%	n=345 46%

Table 4:13 shows that 2f (*getting feedback on students' learning experience* – 93 per cent) and 2e (*to help with paper refinements* – 89 per cent) attracted the highest ‘yes’ response. These high ‘yes’ responses suggest that most teachers claim interest in checking with students about how they are experiencing the teaching and the courses they are undertaking, but also are interested in gaining ideas about fine-tuning their courses to improve students’ experiences.

Simultaneously, 2a (*it is required*) and 2b (*for own professional development*) both attracted similar ‘yes’ responses of 85 per cent and 87 per cent respectively. Not far behind, at 80 per cent, was 2c (*for my promotion application*). Once again, a relatively high ‘yes’ response indicates that there is a sensitivity to the two main purposes for which evaluations have been shown to exist, namely, for accountability as well as for professional development purposes.

Interestingly, 2g (*to provide feedback to my students*) and 2h (*to report on quality matters to relevant internal and external bodies*) were lowest of the ‘yes’ responses (44 per cent and 46 per cent respectively) and highest of the ‘no’ responses (56 per cent and 54 per cent respectively). Sub-question 2f (*to get feedback on my students’ learning experiences*) was the highest ‘yes’ (93 per cent). This suggests that teachers see evaluations as providing feedback to them, *about their students* and for their own use, but *not for their students*. A similar, what might be called ‘teacher-focussed’, view about evaluation may be the reason why there is a sharp contrast between 2e (*to help with paper and course refinements*) – again scoring highly on the ‘yes’ responses (89 per cent) – and 2h (*to report on quality matters*), which scored low on the ‘yes’ responses (46 per cent) and second highest on the ‘no’ responses (54 per cent). These results may indicate that teachers feel they are required to administer evaluations (2a *because they are required*) but the information is primarily for them (2b *own professional development*) and for their teaching (2e *paper refinements*). A similar teacher focus is indicated by other high ‘yes’ scoring items, namely, 2c (*for my promotion application* – 80 per cent) and 2d (*for goal setting/salary review/confirmation* – 66 per cent). Both 2c and 2d point to expectations, even requirements, that are part of the process of demonstrating expertise, effectiveness, capability, development and potential to external bodies such as institutions.

Uses of evaluation data noted by respondents in the comments they provided to Q3 (*If there are any reasons why you use student evaluations/appraisals not covered by the above Q2, please outline them here*) were not very different from the uses that appeared within the Q2 prompts. Of the respondents who commented on Q3 (n = 106), some made general statements that made no reference to the Q2 sub-questions (15 or 14 per cent). The majority (82 or 77 per cent) elaborated on or explained the responses they had already given to Q2. For example:

Appraisals comprise the primary assessment of the content and assessment of taught courses. I use them in the absence of any other more sophisticated or really investigative form of analysis, but as they are currently, they are rather too crude and limited in use. (WU, Senior teaching position, 11–15 years’ tertiary teaching experience, continuing, Humanities.)

Improving the course and my own professional development through student feedback is my primary motivation but providing the students with a safe and formal process to have a say is an excellent conflict management strategy so evaluations also provide this important function. They are also a safety mechanism for lecturers – should we have a bad year or a difficult relationship with a student we can go back to evaluations to demonstrate that the situation isn't typical or if it is typical use the feedback to inform our professional development needs. (OP, Senior teaching position, 0–5 years' tertiary teaching experience, permanent, Humanities.)

It gives me a reflection of the level of information delivery and how it is perceived by the students. This has been very good when I teach a heterogeneous cohort of students with varying background basic knowledge. (OU, lecturer teaching position, 11–15 years' tertiary teaching experience, confirmation path, Health Sciences.)

The remainder of respondents making comments in answer to Q3 (nine per cent) said that they use evaluation/appraisal data to enable consultation with external stakeholders, for applying for jobs, for use in publications and research, or because students expect it.

Further insights into perceptions held about evaluations and why evaluations seem to be regarded as worthwhile can be gained from looking at the responses made to Q4 *When you receive the results from your student evaluations/appraisals, do you....* Question 4 comprised the lead question, followed by seven sub-questions/prompts providing examples of possible activities that teachers may engage in when they receive evaluation/appraisal results. The Q4 sub-questions, 4a to 4g, requested responses on a five-point Likert scale (Always 1 to Never 5). The responses to Q4 highlight connections between perceptions and behaviours.

Table 4:14 presents the total overall distribution across the three institutions of responses to Q4. See also the means and medians in Table 4:15.

It can be seen in Table 4:14 that 4b *read the open comments from the students* is a frequent activity of teachers from across the institutions (95 per cent, 1 and 2). Teachers also 4a *spend time going over the responses* (87 per cent, 1 and 2) and 4g *actively look for feedback about*

teaching and assessment (77 per cent, 1 and 2). In addition, they also claim that they 4d *compare results with previous evaluations/appraisals data* (77 per cent, 1 and 2). On the other hand, they are much less likely to 4c *provide students with feedback* (16 per cent, 1 and 2; 15 per cent, middle ground; 69 per cent, 4 and 5) or 4f *seek assistance with interpreting the results* (12 per cent, 1 and 2; 14 per cent, middle ground; 74 per cent, 4 and 5). In comparison, variation across the five-point Likert scale was generally less marked for 4e *discuss the results with colleagues/teaching team* (47 per cent, 1 and 2; 25 per cent middle ground; 27 per cent, 4 and 5).

Interestingly, while the overall response to 4c *provide students with feedback on results* and 4f *seek assistance with interpreting results from others* suggests a generally low level of activity in these areas, and responses to 4d *compare the data with previous evaluations/appraisals* and 4e *discuss results with colleagues/teaching team*, an average to high level of activity, there are some variations between the frequency of responses among institutions. These variations point to possible differences of institutional context. Where 4c *provide students with feedback on results*, is concerned, OP stands out as it had 55 per cent, 4 and 5. The two universities, on the other hand, attracted 71 per cent (OU) and 70 per cent (WU) of fours and fives. For 4e *discuss the results with colleagues/teaching team* OP responded with a much lower percentage of fours and fives than the two universities (OP, 13 per cent, 4 and 5, WU, 33 per cent and OU, 27 per cent). OP also responded more highly with ones and twos than the universities, the biggest difference being with WU (OP, 64 per cent; WU, 36 per cent; OU, 48 per cent). OP is also different when the results of 4f *seek assistance with interpreting results from others* are examined. Just over half OP respondents said that they are not likely to seek assistance (55 per cent, 4 and 5) compared with three quarters and more at the two universities (OU, 73 per cent and WU, 82 per cent). OP's ones and twos are also higher than the two universities (OP, 25 per cent, 1 and 2; WU, eight per cent, 1 and 2; OU 12 per cent, 1 and 2).

Table 4:14 Overall Responses to Q4 (When you receive the results from your student evaluations/appraisals, do you...)

		1=Always (% 1&2)	(% 3)	5=Never (% 4&5)
4a <i>Spend time going over the data and responses?</i>	Combined institutions	87%	8%	4%
	OU	89%	7%	4%
	WU	81%	14%	5%
	OP	91%	5%	4%
4b <i>Read the open question comments/responses made by the students?</i>	Combined institutions	95%	2%	3%
	OU	96%	2%	2%
	WU	91%	4%	5%
	OP	96%	2%	2%
4c <i>Provide students with feedback on the results?</i>	Combined institutions	16%	15%	69%
	OU	14%	14%	71%
	WU	16%	15%	70%
	OP	26%	19%	55%
4d <i>Compare the data with previous evaluations/appraisals?</i>	Combined institutions	77%	13%	10%
	OU	81%	11%	7%
	WU	74%	13%	13%
	OP	65%	21%	14%
4e <i>Discuss the results with colleagues/teaching team?</i>	Combined institutions	47%	25%	27%
	OU	48%	24%	27%
	WU	36%	30%	33%
	OP	64%	22%	13%
4f <i>Seek assistance with interpreting the results from others?</i>	Combined institutions	12%	14%	74%
	OU	12%	15%	73%
	WU	8%	10%	82%
	OP	25%	20%	55%
4g <i>Actively look for feedback about teaching and assessment?</i>	Combined institutions	77%	13%	11%
	OU	75%	13%	12%
	WU	76%	14%	10%
	OP	84%	9%	7%

Table 4:15 shows that all items on Q4 give a significant difference between two or more institutions. Median tests show there are differences between two or more groups for each sub-question for Q4. Fairly similar for all, 4c *provide students with feedback on the results* and 4f *seek assistance with interpreting the results from others* are least done by all.

Table 4:15: Q4 Median Test Results

Q4	<i>a. Spend time going over the data and responses</i>	<i>b. Read the open question responses made by students</i>	<i>c. Provide students with feedback on the results</i>	<i>d. Compare the data with previous appraisals</i>	<i>e. Discuss the results with colleagues/teaching team</i>	<i>f. Seek assistance with interpreting the results from others</i>	<i>g. Actively look for feedback</i>
OU	1.46	1.16	4.01	1.76	2.72	4.06	1.88
WU	1.67	1.36	4.00	2.01	2.97	4.31	1.81
OP	1.34	1.17	3.55	2.22	2.28	3.52	1.55
Mean	F=7.659 p=0.001 (WU & OU)	F=7.975 p=0.000 (WU with OU & OP)	F=7.669 p=0.000 (OP with OU & WU)	F=11.817 p=0.000 WU⇔OU⇔O P	F=10.110 p=0.001 WU⇔OU⇔O P	F=15.971 p=0.001 WU⇔OU⇔O P	F=4.2200 p=0.015 OU & OP
Independent Median Test	p=0.000	p=0.000	p=0.001	p=0.000	p=0.001	p=0.000	p=0.003
Grand Median	1.00	2.00	4.00	2.00	3.00	4.00	1.00

4.3.1.2 Teachers' perceptions of student evaluations – probing beliefs and emotions through the interviews

The interviews aimed to probe the beliefs, values and possible emotions that might influence the perceptions about evaluations that were uncovered in the questionnaire data. The analysis of the interviews for research question 1 focussed on respondents' views on three broader dimensions that the literature indicated may usefully shed light on lecturers' perceptions of the value of students' feedback in student evaluations. These areas were lecturers' beliefs about the nature of the teaching and learning process, attitudes to students' competence to make judgements and lecturers' personal emotions. While the whole interview data set was considered, particular focus was placed on responses to question 2a *how do you see yourself as an educator?*, question 2b *can students make judgements about the quality of the teaching and their learning experience?* and questions 2c and 2d, *when you receive the results of evaluations from students, how do you feel? and why?* The investigation also aimed to see possible correspondences and connections between these three points. The feedback around these points is grouped by institution.

4.3.1.2.1 University of Otago

a) *Interview question 2a: How do you see yourself as an educator?*

In the comments about their teaching conceptions, there were no particularly noteworthy patterns in the responses of OU participants. Reference to developed educational theory was absent, and the self-conceptions that were articulated were very varied. These conceptions ranged from those who focussed on the activities of the teacher in the process of delivering content, to those who talked more about the students, but not necessarily in terms of learning outcomes.

Seven participants talked about their role primarily in terms of teacher behaviours. These comments on perceptions of desirable teacher behaviours were in some instances linked with the conception that teaching is about the delivery of content. Comments include:

Teaching involves a well-taught basis of factual knowledge.

Communication, interest, fun.

Teaching is a basic thing. Get people's interest and try to explain it to them.

Communication, engagement. Many students don't like the subject. Make it fun.

Keep students' attention.

Tries to renew keep fresh, change.

Eight OU interviewees saw their goals for teaching primarily in terms of their students. The level of pedagogical awareness varied: some articulated a well-defined vision of their goals for their students, while others simply acknowledged that their students ought to be the primary focus of their work. Examples of comments include:

Ask them questions, get them to try and think about it for themselves.

Challenges students to learn independently and engage in critical reflection.

Sets students up to conduct their own research. Setting up students for the future and the reality of their jobs is important.

Facilitator. Maybe some scaffolding. Bringing the students to the subject rather than that's how it is.

Engaging students in something and helping them learn how to do that something.

My priorities are to improve the learning experience for the students, but that's a job that's never done really.

Research was mentioned by a number of the interview participants (eight). Some of these respondents conceptualised themselves primarily as researchers. For example:

Tries to renew, keep fresh, change. Researcher first.

Teaching is not primary focus. I wish to be an excellent researcher but am content to be good at teaching.

Another theme was the need to balance teaching and research or the tension related to dividing one's time between them. For example, one interviewee believed that the wish to focus on teaching is under pressure because of an institutional focus on research:

Teaching is a priority, but I'm under pressure to spend more time on research.

Another factor that was mentioned as impinging on teaching beliefs is the interface between clinical or applied learning and academic and theoretical knowledge.

b) Interview question 2b: Can students make judgements about the quality of the teaching and their learning experience?

The findings about academics' views about students' capacity to make judgements were very mixed. Only three OU interviewees unequivocally supported the view that students could make these judgements, while another two were mainly positive with minor reservations. Examples of comments by those who strongly endorsed the notion that students could make judgements include:

Yes, students have many years of knowing what is good teaching or a good course.

Yes. And I'm aware when teaching and taking into account what they are doing and how they are responding.

An example of a primarily positive view is:

To a point they can, as they have experience of the university environment, so they can comment on programmes and performance.

The largest number of participant views fall into a middle group, ranging from what can be termed positively disposed with reservations (six) to a degree of scepticism (six).

The positively ambivalent group made comments such as:

Yes, but not necessarily informed judgements. If it's too hard they may judge negatively.

Yes, can make useful judgements, but not able to help with the fine tuning.

A degree of scepticism is evident in comments such as:

Yes, but you should take it with a pinch of salt.

Students are influenced by many things and just want to be given the answers.

Three interviewees expressed emphatic views that students could not make judgements, in comments of this nature:

No, I know many of them would know if they found you likable. It's a popularity contest a lot of it.

There was some correlation between teaching conceptions and attitudes to students' capacity to make judgements, for example, all three participants who were emphatically positive about students' capacity to make judgements, highlighted students and their learning in their articulation of their beliefs. One of these interviewees spoke about teaching as:

Bringing the students to the subject, rather than that's how it is.

Likewise, those who were mainly positive about students' capacity to judge (two) expressed student-centred views of teaching such as:

My priorities are to improve the learning experience, really for all our students.

A similar trend was observable in the views of those who were positively disposed with reservations. Four out of six in this group had a definite focus on the students in their belief statements. Examples include:

Engaging the students in something and helping them learn how to do that something.

Setting students up for the future.

The number of those who focussed on students in the belief statements declines with the more sceptical respondents (two out of six) with one being categorically opposed to student-centred learning, commenting that:

It's not about improving practice or student-centred folk pedagogy.

For the three respondents who expressly rejected the idea that students can judge the teaching, there was no discernible relationship with their belief statements. Two of these did mention students, but there was no sense of what they wanted for their students in terms of learning outcomes.

c) Interview question 2c: When you receive the results of evaluations from students, how do you feel? and 2d, Why?

Among those interviewees who commented on their emotional reactions to student evaluations (10) there was a mix of positive and negative sentiments. The language used by interviewees hints at the emotional dimensions of receiving student evaluations. Even those who expressed positive emotions often used non-neutral language indicating feelings of emotional intensity in relation to student evaluations. For example, one interviewee remarked that he is “excited, cleansed” on receipt of positive evaluations, but “pissed off” if they are negative. The use of emotionally charged language was intensified when negative feelings were reported. For example, one participant referred to “fear” and “feeling vulnerable” and suggested that evaluation results were “sensitive”. Another participant said that if results were bad, they would be “crushed”, while yet another used the phrase “can feel down”. Still others spoke of “anxiety” or being “angry”. Two of the interviewees who expressed negative emotions linked these sentiments to the institutional use of evaluations.

There was no obvious correlation between negative emotions around evaluations and particular teaching and learning beliefs.

4.3.1.2.2 The University of Waikato

a) Interview question 2a: How do you see yourself as an educator?

A strong theme in the group of responses to question 2a of the interviews is a focus on handing something over to the students. Nine out of the twenty interviewees, to different degrees, talked about their practice in terms of a transmission paradigm. However, there was

considerable range within this group. The variations include those who spoke of a relatively clear cut process of knowledge and skills transfer, those who see the teacher's role as helping to mediate, organise and interpret the material for the students, and others who emphasised application of the learning materials in practical contexts. Others stressed the importance of the students connecting with, and making sense of, the materials. Examples of comments include:

To teach the knowledge and the principles to the students for the curriculum, the course content.

I come from a background in which lecturers impart information and students aren't expected to respond to it [but acknowledges that this is changing because of questioning students].

Likes to communicate things clearly. Ensures no ambiguity. Real life connections to abstract concepts.

I give them access to assistance and clarify concepts for them.

I bring my enthusiasm for the topic to the class. Because of my background in research, I try to bring as many practical examples as possible.

It's that whole philosophy of trying to make it interesting by using practical examples.

I'm very keen to get students interested and involved.

To inspire students to understand the material.

To foster understanding, not just absorbing information.

By contrast, with this strong focus on handing over material or skills to students, a number of WU interviewees (six) discussed teaching and learning in terms of developing students. A number of these comments were also linked to the perception that educators have a social responsibility. Examples of these comments include:

I like to think I have a lot to offer in the way of opening people's minds; encouraging to think for themselves. Each student is a potential centre for improvement in society.

At a macro-level, my priorities would be to contribute to the education and personal growth of citizens in this country.

I work closely with students to encourage them to push the boundaries.

I want them to think about issues more deeply, not just rote learn.

Encouraging students to ask questions and become suspicious researchers that question the taken for granted. Lifelong learners.

I think I'm very keen to enable students to see a bigger picture... I'm also concerned about social justice and hope that students will come to see that as important in their practical work.

Other conceptions of teaching and learning that were voiced by WU interviewees include the idea of the teacher as mentor (two), the view that learning is fostered through students' activities (one), the notion of teaching as a 'subset of research' (one) and the conception that students and teachers are co-learners (one).

b) Interview question 2b: Can students make judgements about the quality of the teaching and their learning experience?

Findings from WU demonstrate a relatively positive disposition about students' capacity to make judgements on teaching and their learning experiences. Half of the interviewee group (10) expressed support in varying degrees for the idea that students could make pertinent judgements about the quality of their learning experiences. There was a noticeable correspondence between those interviewees who articulated conceptions of teaching with a strong focus on learner development and enthusiastic support for students' ability to make judgements. Five of the ten interviewees who expressed developmental views of education were positive about students' capacity to judge. Interestingly, one interviewee who sees the teacher's role as enabling students to acquire the tools to address social injustice, did not fit this pattern. That interviewee was more ambivalent about the judgements that students make, because, from the interviewee's perspective, the views expressed by the students had been influenced by a value-for-money orientation to education as a consequence of having to pay fees.

Examples of positive views of students' judgement capabilities include:

I think students are very tolerant of less than perfection, and they are reluctant to be critical, but if things are not good they will make their feelings known. But it takes quite a lot to provoke them.

Yes. All consumers are. Not that they're just consumers...Some academics think they're not... I think that's insulting. [conception of self as educator: I think I have a lot to offer in the way of opening people's minds]

Yes they can. They can seriously make judgements.

Too damn right. I think they are eminently well placed. [conception of self as educator: to contribute to the educational and personal growth of citizens]

Yes, they're pretty astute really. [conception of self as educator: I want them to think about issues more deeply, not just rote learn]

Most definitely. Students at the tertiary level know what they want to achieve in order to progress.

Seven WU interviewees gave some credence to students' ability to make judgements about their learning experiences, but expressed a range of reservations or uncertainty about the quality and value of students' judgements. These reservations varied from minor to more serious concerns.

Examples of comments include:

Totally capable of evaluating, but subjective and necessarily partial. It's always partial, because they simply don't have the wide view often.

They now evaluate according to the value they get because students have to pay. It's made them critical.

I believe there is a cultural bias. If I were a student I would not want in any way to embarrass the person who has been teaching me [for Māori and Pacific students].

Not first years.

I just can't trust it there are too many variables.

Of this group who expressed some uncertainties about students' competence to judge, five articulated some form of transmission conception of teaching and learning, while one spoke of being a mentor and one saw teaching as collaboration with students.

Three WU interviewees expressed strong convictions that students could not make judgements about the quality of their learning experience. Their comments include:

I used to believe that, but now I no longer believe that. I think in terms of how...students are believing they are buying a qualification...it's more like purchasing their degree. It has a significant impact on their judgement of what they're learning. [conception of self as educator: I'm only really part of the equation; other parts include students, facilities, the mixture of how students interact, the physical classroom].

No I often feel disappointed. When I read them I think now what did I do to you again? [conception of self as educator: transmission plus practical application]

Not undergraduates. [conception of self as educator: teaching as a subset of research].

Each of these interviewees who were negative about students' ability to judge seemed to hold a different teaching conception, but it is noticeable that none of them expressed conceptions that focussed significantly on student development or social awareness agendas.

c) *Interview question 2c: When you receive the results of evaluations from students, how do you feel? and 2d: Why?*

Four interviewees used language that suggested they attach emotional significance to evaluations/appraisals and these emotions can be negative. Responses include:

Disappointed, because they invariably give me great marks, but they don't give me anything positive on which to work.

I was terrified in my first year.

I rip open the packet or download it.

Emotionally, I'm pretty disappointed if I get a score of 2 or more.

You can be a good teacher, but if they haven't understood it, that can be frustrating.

Three other WU interviewees referred more directly to a negative impact, which they experienced on a personal level. Their comments were:

I often feel disappointed when I read them. I think, 'Now what did I do to you again?'

I've been very wary of asking for personal views, because if you take those on board they can be used against you.

I must say I have a tendency to read the comments and focus more on the negative comments than the positive. The one that was pretty negative, I felt deeply distressed by it.

Two of these WU respondents who indicated personal negative emotions, made comments that suggested transmission conceptions of teaching and they believed that students could not make judgements about the quality of their learning experience. The third, who acknowledged a tendency to focus on negative comments, held a developmental teaching philosophy and believed that students are competent to judge the teaching and their learning experiences.

4.3.1.2.3 Otago Polytechnic

a) *Interview question 2a: How do you see yourself as an educator?*

The strongest theme to emerge in the interviews with OP participants in relation to this question was a strong focus on students and their learning outcomes. Many interviewees placed the students in the foreground of their articulation of their conceptions of themselves as educators. Twelve interviewees directly referenced the students in their responses. Correspondingly, there was not such a strong emphasis on the delivery of content in the comments of OP educators. Examples of the numerous student-focussed views include:

I'm there for them to facilitate their learning. Students are core.

100% available for students. Helping students develop life and vocational skills.

As a facilitator. Students and teachers facilitate each other's learning.

Associated with this student focus, there was a marked emphasis on the educator's role in preparing the students for the workforce, skills development and long-term learning (11). OP

educators appeared to be very conscious of their responsibilities to, and for, students beyond the confines of the classroom. Examples of comments include:

Helping students develop life and vocational skills.

Help students to see the relevance of learning, transferability.

I want students to be knowledgeable reflective practitioners in their field.

b) *Interview question 2b: Can students make judgements about the quality of the teaching and their learning experience?*

In response to question 2b of the interview, the findings do not correspond well with the strong emphasis on students in the OP educators' beliefs statements attained through question 2a. The view of the importance of the students in the teaching and learning process was not matched by a noticeably strong perception that students are competent, perceptive judges of the quality of teaching and their learning experiences. While some interviewees from OP were emphatic that students were more than capable of making these judgements, the bulk of the views sat in the ambivalent middle ground (13 respondents expressed some level of ambivalence). No interviewees rejected outright the idea that students could judge the teaching and learning. Examples of comments include:

To some extent, but they don't always know what they need to learn.

Can describe a lecturer. May not be able to make a judgement on what is good teaching or not.

Students are good at judging whether they like it or not, but not the quality of the teaching.

c) *Interview question 2c When you receive the results of evaluations from students, how do you feel? and 2d, Why?*

A small but emphatic group of interview participants from OP used language and expressed views that indicated a powerful element of emotional sensitivity around evaluations when answering questions 2c and 2d. Six openly discussed their feelings of vulnerability in relation to negative comments and their feelings of discomfort when receiving criticism. Examples of comments include:

I'm always scared that I'll see something I don't want to.

Students have bullied staff and they use evaluations as an opportunity to dump on staff.

For me, I take it a bit personally. I suppose that if a student doesn't pass then they take it out on you.

I will gravitate towards the negative comments and avoid the compliments.

4.3.2 Conclusions

In summary, while responses to Q17 of the questionnaire focussed on teachers' views on the personal worth of evaluations/appraisals, Q2 drew attention to how teachers' behaviours may reflect those views of personal worth. Overall, Q17 provides evidence to show that there is a positive disposition toward evaluations and the comments provided in Q18 indicated that the meanings attributed to 'worth' were related to factors that limited and contributed to that sense of worth. Additionally, while responses to Q2 reported why teachers use student evaluations/appraisals, Q4 responses provided some sense of how teachers' behaviours reflect those Q2 reasons for using evaluations and Q17/18 views about personal worth of student evaluations/appraisals. Furthermore, responses to Q17/18, which illustrated two major themes – (1) factors contributing to teachers' sense of worth of evaluations; and (2) factors limiting teachers' sense of worth of student evaluations – highlighted the variety of motivational reasons behind uses of evaluation/appraisal data. These motivations span a continuum between those that tend to be intrinsic, for example, using student evaluations for professional development, for better teaching, for students, and those that tend to be more to do with external imperatives, including requirements of the job, for reporting quality, to move/survive/progress through the system. Comments from the interview data added some elaboration to factors as they are perceived by teachers, particularly around the question of whether students are able to make judgements about the quality of teaching and their learning experiences and the personal responses to receiving feedback from students. These two sets of motivations are not necessarily unrelated. Motivation to move through the system or the desire to advance or survive may have intrinsic origins, but the process of moving through the system means that extrinsic requirements must be satisfied. Along the way, the emotional response to critique received from students plays a part in influencing the value placed upon

the worth of student comment and perhaps even the worth of the student evaluation system itself.

The desire to do a good job as a teacher and to do the best for the students is a strong motivator, so gathering evaluation data to improve teaching and courses is a high priority. The interview data showed that there is some relationship between the conceptions held about teaching and learning and the worth that teachers place on student feedback. Educators expressing student-centred conceptions tended to place a higher value on student involvement in evaluation processes than those expressing teacher-centred/transmission views of teaching. This desire to be an effective teacher is often accompanied by an essential need to be accountable to oneself, to ones' students and to one's institution, and this desire may be driven externally and/or internally. There is accountability to external bodies (essentially to let *others* know that teaching and courses are effective) and this involves meeting requirements and going through processes in order to 'get on' or to promote oneself, to cope within, or make one's way through, the system. Alongside this, there is accountability to oneself, which also might include accountability to one's students and the discipline or profession. Essentially, this accountability is to satisfy oneself that teaching and courses are effective and involves gathering evaluation/appraisal data in order that the teacher knows what is going on and can learn from and refine/develop/improve their teaching and subsequently student learning.

4.3.3 Question 2: What factors (causes, influences) affect these views?

The culture of New Zealand tertiary institutions is diverse and influenced by a number of factors, but a starting point to understanding the diversity is the influence of the national Tertiary Education Strategy 2010–2015 (<http://www.minedu.govt.nz/NZEducation/EducationPolicies/TertiaryEducation/PolicyAndStrategy/TertiaryEducationStrategy.aspx>). Excerpts from Part Three of the Strategy were outlined in section 2.2.1. The different emphasis on research and level of education between the polytechnics and the universities is explicit in the Strategy. Universities are expected to be research-led with strong graduate and postgraduate education and be recognised at an international level. Polytechnics are expected to provide education with a vocational focus

and assist student progression to higher levels of learning. This is to be supported and enhanced by applied research.

This difference is also reflected in the vision, mission and values of each of the institutions involved in this study. In Table 4:14 it can be seen that OP is highly focussed on learner experience and the provision of opportunities for learning. The universities, on the other hand, focus on research and international excellence.

Based on the expectations of educational providers as stated in the Strategy and the claims the participating institutions have made in their vision, mission and values statements, it was expected, in turn, that a stronger focus on teaching in OP than in OU and WU would be detected as the data being gathered during this study were analysed. These expectations are now explored in this section of the report.

4.3.3.1 Institutional requirements

As described in section 4.3.1, one of the reasons many staff use student evaluations/appraisals is because they are required to by their institution. Question two in the questionnaire asked respondents to identify why they use student evaluations/appraisals (*Please identify why you use student evaluations/appraisals.*), with one of the possible selections being, 2a *because it is required by my department/school/institution*. The percentage of respondents who answered ‘yes’ to Q2a is shown in Table 4:16, below.

Table 4:16 Percentage of ‘Yes’ Responses to Q2a (because it is required by my department/school/institution)

Institution	% “Yes” responses to Q2a
OU	82%
WU	85%
OP	97%
ALL	85%

The distribution of responses shown in Table 4:16 shows a similar percentage for the two universities (82–85 per cent), but OP is much higher (97 per cent). A Pearson Chi-Square test was undertaken to see if there was a significant difference in the proportion of ‘yes’ and ‘no’ responses from each institution. For OP the proportion of ‘yes’ responses was significantly higher ($p = 0.000$). This suggests that a relatively higher proportion of staff at OP perceived

that they need to run student evaluations/appraisals to meet institutional requirements than staff at OU and WU, although the differences between the institutions were not significantly tested. Possible reasons for this may be that OP has more rules in place about the frequency of evaluation. As outlined in Table 2:5, at OP, the expectation is that staff will seek feedback at least annually and that courses are evaluated at least once a year. At WU there is a mandate to evaluate using the central evaluation system, but less frequently, once every two years. At OU there is no explicit requirement to evaluate teaching or courses, unless applying for promotion or on confirmation path. In this case, it is expected that teachers will evaluate all their teaching over a three-year period for promotion, and more often for confirmation path staff. In summary, though, overall, the percentage of 'yes' responses is high, indicating that institution requirements are a main driver for student evaluations at all three institutions, and so should be considered an important influencing factor for staff.

Further analysis of Q2a was conducted to search for statistically significant differences within the following three demographic groups: years of tertiary teaching experience (Q26), nature of employment (Q27), and current position (Q28), for each separate institution. The relationship between categorical variables was analysed using the Pearson Chi-Square test.

For the three institutions, there were no significant differences in the proportion of respondents answering 'yes' compared with those answering 'no' by years of teaching experience in a tertiary institution (Q26). OP has a high proportion of teachers with 0–5 years' teaching experience (27 per cent within the category) stating student evaluation is a requirement, compared with OU (15 per cent within the category) and WU (12 per cent), however the results are non-significant.

At OU, the proportion of 'yes' responses to Q2a differs by current position (Q28) and nature of employment (Q27) (Pearson Chi-Square $p \leq 0.05$). The 'lecturer teaching positions' together with 'senior teaching positions' compared with 'junior teaching positions' (see categorisation described in section 3.4.3.1.1), and permanent staff compared with confirmation staff (see Table 4:6), have a significantly higher proportion stating they use student evaluations as a requirement. In other words, at OU, staff who are in the 'lecturer teaching position' and 'senior teaching position' groups are more likely to use student evaluations because it is a requirement, as opposed to those in the 'junior teaching position' group who are less likely to use student evaluations as a requirement. This difference does

not exist at OP or WU, as the data showed no proportional differences by position and nature of employment for those two institutions.

4.3.3.2 Other institutional influences

Section B of the questionnaire asked for more depth about respondents’ perceptions of evaluation/appraisal data and connections with teaching practices. Influences on how teaching decisions are related to perceptions of institutional use of the data was explored in Q11 (*To what extent does your institution’s use of student evaluation/appraisal data influence your teaching decisions?*) with opportunity for participants to make free text comment in Q12 (*Please explain your answer to Q11*). Respondents were asked to answer Q11 using a five-rating Likert scale with a descriptor at each end: 1=A great deal and 5=Not at all.

The distributions from Q11 appear in Table 4:17, below, and show the combined totals for ratings 1 and 2, and 4 and 5, as well as the ‘middle ground’ (rating 3). The distributions show a difference in perception between the three institutions, with the most marked difference between the two universities and OP. Over half of OP staff rated 1 or 2, while just over a third of OU staff and a quarter of WU staff rated 1 or 2.

Table 4:17 Combined Total Ratings for Q11 (To what extent does your institution’s use of student evaluation/appraisal data influence your teaching decisions?)

	1&2 %	middle ground %	4&5 %
OU	35%	30%	35%
WU	25%	29%	46%
OP	53%	27%	20%
All	35%	29%	36%

The ANOVA also indicated that OP has a significantly lower mean and median than the two universities. (ANOVA p=0.044; independent samples median test, p=0.0139). This is shown in Figure 4:1 below. The lower the mean indicates the greater the level of institutional influence.

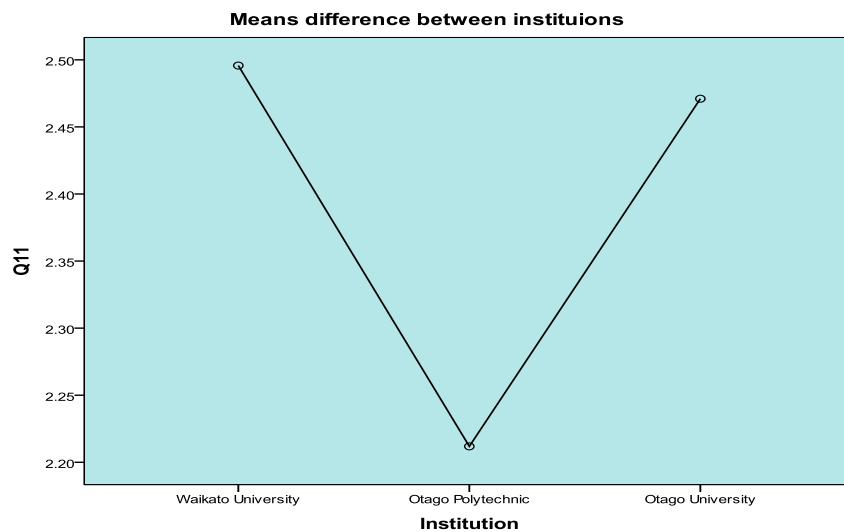


Figure 4:1 Perceptions of institutional influence on teaching decisions across institutions. The results shown in Figure 4:1 and the frequencies shown in Table 4:17 indicate that OP staff claim to be more influenced in their teaching decisions by institutional use of student evaluation data than either OU or WU staff.

Differences are also evident when Q11 responses are examined against the length of tertiary teaching experience demographic (Q26). The means test for each institution for years of tertiary teaching experience showed no significant difference between the groups (0–5 years; 6–10 years; 11–15 years; 16–20 years; and 21+ years) from within each institution. However, what is interesting is when the means across the three institutions are compared. Table 4:18 shows the group with the lowest means for each institution (the lower the mean indicating the greater the level of institutional influence).

Table 4:18 Significant Differences Between Tertiary Teaching Experience (Q26) and Q11 Responses

Institution	Years of tertiary teaching experience with lowest mean (i.e., most influenced by institutional use)
OU	11–15 years
WU	0–5 years
OP	16–20 years

Table 4:18 shows that within each institution, compared with other groups, for OU, the 11–15 years’ teaching experience group, for WU, the 0–5 years’ group and for OP, the 16–20 years’ group are more likely to be influenced by institutional use.

While it is difficult to identify why the 0–5 years of tertiary teaching experience group at WU should be more influenced by the institutional use of appraisal/evaluation data, possible reasons that might account for the connection for the OU and OP groups can be suggested. At OU, respondents belonging to the 11–15 years of teaching experience group are more likely than other groups to be applying for significant promotion, and, presumably, would be more sensitive to the promotion requirements of the institution than perhaps the other groups. A possible reason to explain OP staff with 16–20 years of teaching experience being more influenced by institutional use may be that many within this service band hold senior positions, such as principal lecturer, and will also be programme managers and/or course coordinators. They will be, therefore, more closely involved in the management of the academic area they are in, more aware of the need to collect data from students and then how those data are used to report on the performance of the School.

The ratings data above provide a broad picture of perceptions of institutional use of student evaluation/appraisal data having influence on teaching decisions. To gain a better understanding of why these perceptions exist, and the issues surrounding why, the qualitative data collected in Q12 (*Please explain your answer to Q11*) were examined.

In Appendix 7, note that of all the comments, those from OP staff made up substantially fewer than from WU and OU (OP 25 per cent; OU 64 per cent; and WU 76 per cent). A detailed description and breakdown of themes that emerged from the comments to Q12 supplied by respondents can be found in Appendix 7. Mostly, the comments pointed to use for promotion and quality control aspects. These comments are grouped around sub-themes 1 and 4 (specifically, 4b and 4c). Of note, university staff made over twice the percentage of comments about the use of evaluations for promotion purposes (29 per cent OU and 30 per cent WU) compared with OP staff (12 per cent). Some examples of these comments include:

Student evaluations are a required component of salary and promotion review, but I would do them anyway. (Q12, sub-theme 1a, OP, senior teaching position, 21+ years' tertiary teaching experience, permanent position, Sciences.)

Because of the need to have positive student evaluations for promotion/confirmation etc.; there is a tendency to try to do what the students want (which is often to be spoon fed) or to choose questions that highlight your

strengths. (Q12, sub-theme 1b, OU, senior teaching position, 21+ years' tertiary teaching experience, permanent position, Sciences.)

The institution does look at the appraisal data and it has influenced my career development from the probationary period, through to promotion to senior lecturer. Gaining my annual increments of salary is partially dependent on the appraisal results. So, there is a need to keep these looking good. However, I received perfect ratings numerically in several papers during the past couple of years (1.0 for teaching), and the institution did not recognise this as particularly significant. It is easy for those in positions of responsibility to dismiss these results as being less significant due to smaller class sizes, or to suggest that I should be putting more time and effort into research and publications, rather than teaching. I have been advised by senior colleagues that any score better than a 1.6 represents wasted effort that should be directed elsewhere. As long as the scores are under 2, few questions are asked. There is no push for excellence and little celebration of this, except for the few who win excellence awards in the formal sense. (Q12, sub-theme 1c, WU, lecturer teaching position, 6–10 years' tertiary teaching experience, continuing position, Humanities.)

I'm aware that my teaching is scrutinised in this indirect way, but I choose not to worry about it. (Q12, sub-theme 4b, WU, lecturer teaching position, 11–15 years' tertiary teaching experience, continuing position Humanities.)

Appendix 7 also shows that nine per cent overall of comments made in answer to Q12 (35 comments from OU, 28 from WU and four from OP) indicated no awareness of how the institution uses the data (sub-theme 4a). For example,

Don't know if they do use the evaluations and even if they did; can't see what use they get out of them. (Q12, sub-theme 4a, OU, lecturer teaching position, 21+ years' tertiary teaching experience, permanent position, Humanities.)

I didn't think the institution used the data, I thought it was confidential to the staff involved in the course. (Q12, sub-theme 4a, WU, senior teaching position, 21+ years' tertiary teaching experience, continuing position, Sciences.)

Wasn't really aware that Polytech does use it beyond me using it as evidence to support a case. Tend to assume it's pretty private otherwise. (Q12, sub-theme 4a, OP, lecturer teaching position, 21+ years' tertiary teaching experience, permanent position, Humanities.)

Of those who identified the use of evaluation/appraisal data for development purposes (overall 20 per cent), OP staff were the highest at 31 per cent, followed by OU at 22 per cent and WU at 16 per cent. (Note that the OP percentages are distorted somewhat by the low response rate to the question: the 31 per cent equates to only 13 comments under this sub-theme). The quotes below from sub-theme 3a provide examples of how respondents expressed the connection they perceived to exist between evaluation/appraisal data and its use for developmental purposes.

I use them primarily for deciding how to change and modify my teaching practices. I used student evaluations well before the OU decided to institutionalize them but from the start up until now my opinion is that the prime purpose is to help inform the teacher how the teaching is going and ways to change and improve. (Q12, sub-theme 3a, OU, senior teaching position, 21+ years' tertiary teaching experience, permanent position, Sciences.)

The major reason I carry out appraisals is to gain feedback about, and attempt to improve, my teaching. All other needs for these data are secondary as far as I'm concerned. (Q12, sub-theme 3a, WU, lecturer teaching position, 16–20 years' tertiary teaching experience, continuing position, Sciences.)

I have never thought about the use of evaluations influencing what I do but...I think if I started planning courses based on the fact that they will be evaluated it would be a sad day signifying loss of passion and enthusiasm for what I do. It would be a sad day for the polytechnic too – evaluations shouldn't be a punitive measure – they are too important as a professional development tool. The influence evaluations have on my teaching and planning is in the questionnaires themselves and the question pool! This is an extensive range of questions that really make me stop and think, am I doing what I'm supposed to do? It is interesting to think that evaluations have a direct influence on professional

development even before the feedback comes in! (Q12, sub-theme 3a, OP, senior teaching position, 0–5 years' tertiary teaching experience, permanent position, Humanities.)

The other broad theme (Theme 3) covered general concerns with the evaluation instruments (14 per cent overall). These concerns were mainly grouped under sub-theme 3a (8 per cent), highlighting views that evaluation/appraisal data were too 'blunt' or of limited use, and/or critical of student ability to give good feedback or being too influenced by popularity and bias. These views were shared fairly evenly by all three institutions. The following comments provide illustration.

Students at this university generally do not like hard or challenging work. If they are given such; they will often 'punish' the deliverer with negative evaluations regardless of the quality of delivery. (Q12, sub-theme 3a, OU, lecturer teaching position, 16–20 years' tertiary teaching experience, permanent position, Commerce.)

It is difficult to know the influence appraisals have on superiors. It is a tool open to manipulation. I have had some very high and successful student appraisals but little acknowledgement has ever been made of them. Certainly nothing practical has resulted. It is the opinions and comments of students after they have been out working for a few years that I most value. These can vary from those opinions expressed during the course. (Q12, sub-theme 3a, WU, lecturer teaching position, 21+ years' tertiary teaching experience, continuing position, Commerce.)

Again, the surveys do not provide substantial enough feedback to fuel major changes in curriculum and delivery. (Q12, sub-theme 3a, OP, lecturer, 6–10 years' tertiary teaching experience, permanent position, Sciences.)

In summary, the comments revealed a broad range of attitudes towards evaluation use by the institution. Although the developmental aspects were the single highest grouping of comments, the other groupings were quite revealing about the areas of tension staff experience regarding student evaluation, particularly in relation to judgemental/quality

control uses. The number of comments (almost 200) is high enough for these tensions to need further exploration.

4.3.3.3 Internal/external reporting requirements

Another influencing factor examined in the data was the requirement to report effectiveness or quality of teaching and courses to bodies other than those immediately involved in the teaching. To do this, once again responses to Q2 (*Please identify why you use student/evaluations/appraisals*), paying particular attention to 2h (*To report on quality matters to relevant internal and external bodies*) were examined specifically.

The overall responses to Q2h are shown in Table 4:19, below.

Table 4:19 Percentage of ‘Yes’ Responses to Q2h (to report on quality matters to relevant internal and external bodies)

Institution	“Yes” responses (%)
OU	38%
WU	47%
OP	74%
ALL	46%

Table 4:19 shows that 74 per cent of OP staff reported use of evaluations to report to internal and external bodies, compared with 47 per cent for WU and 38 per cent for OU. Two possible reasons may be contributing to this expressed perception of the use of evaluations/appraisals. First, the nature of many of the programmes taught at OP, including all Health Science programmes, have reporting requirements to external bodies such as the New Zealand Nursing Council, and the New Zealand Council of Midwives. Also, all Schools report to external advisory committees. In addition, many of the certificate and diploma programmes offered by OP are unit standard based and have external moderation and reporting requirements. Second, the size and structure of the three institutions may be an influence. The smaller size and flatter organisational structure at OP may mean that staff within its Schools tend to be more directly involved in reporting to external and internal bodies than many staff in the two universities. As a result, OP staff are often more personally involved with the various reporting activities that have to occur, and understand and experience the outcomes and consequences of both external and internal requirements of this type.

As illustration, the following quotes from Q12 of the questionnaire, which asked for comments about whether and how institutional use of evaluation/appraisal data influences staff teaching decisions, highlighted the connection between student evaluation and reporting on quality matters to external bodies. It should be noted that only a small number of staff (four in total) commented directly on external/industry monitoring purposes of evaluation.

Student evaluations can be used in the monitoring process we have from the New Zealand Teachers Council. Feedback may result in changes to assessment. (Q12, sub-theme 1d, OU, junior teaching position, 21+ years' tertiary teaching experience, permanent position, Humanities.)

Results need to be made public so they can be compared, benchmarked, etc. Public results will identify colleagues who are worthy of praise and emulation, and those who need to allocate more effort or work on their teaching. (Q12, sub-theme 1d, WU, senior teaching position, 16–20 years' tertiary teaching experience, continuing position, Sciences.)

It is on the landscape – but the context is far wider than just student feedback. Industry is a huge influence and often the fact that students may find an aspect of a course difficult to stressful is not a good justification for change – often industry which we are accountable to also) directs practice (and it should in my discipline). Research/evidence is also important – so student opinion and experience is one factor but not exclusive. We have to take a responsible approach – and public safety is important to us. (Q12, sub-theme 1d, OP, lecturer teaching position, 6–10 years' tertiary teaching experience, permanent position, Health Sciences.)

The internal/external reporting reason was the second lowest one used by staff as to why they use student evaluations, but the percentages only show some of the picture. As already noted, three quarters of OP staff say they use student evaluations to report on internal and external quality matters. In comparison, less than half of the two universities' staff identified this as a reason. However, when the actual numbers of staff (172 for OU and 92 for WU) are considered, it does show that reasonably large numbers of staff do use student evaluations for

this reason. It may not be their main use, but this type of reporting should still be considered a definite influencing factor when considering the breadth of use of student evaluation data.

4.3.3.4 Institutional culture and expectations

The material presented in the literature review pointed to a connection between institutional expectations, whether real or perceived, teachers' perceptions of evaluations/appraisals, and the teaching decisions they make. From the data collected in this study, it is clear that staff perceptions of institutional use of evaluation data are mixed. The use of evaluations to judge quality through processes such as promotion, does not sit well with some staff, for example:

Institutions disproportionately use appraisal data to evaluate the quality of teaching. To my mind, this is a significant issue in tertiary education, as it perpetuates the transformation of students into consumers. (Q12, WU, senior teaching position, 16–20 years' tertiary teaching experience, continuing position, Commerce.)

While others say they support this use or are not influenced by it, for example:

I think the emphasis is on us to review our own practice in the light of this feedback. I do not feel 'pressured' in any direction by institutional use of these evaluations. So long as learner feedback is gathered, collated and represented coherently by OP systems I can make use of it. (Q12, OP, lecturer teaching position, 0–5 years' tertiary teaching experience, permanent, no division response.)

Each institution has different processes in relation to career progression and promotion (see Table 2:5), which may account for the differences in perception. OP staff indicated high awareness of institutional use of evaluation data, but simultaneously claimed that personal development is the primary purpose of evaluation. It is interesting to note that while the evidence is not conclusive (the OP percentage was higher than the universities', but there were only 13 comments overall, as the response rate to this question was low), appreciation or awareness of the two seemingly opposing purposes can be held by staff at the same time. In other words, for a small number of participants, evaluation for institutional quality control purposes may be viewed quite comfortably as operating alongside evaluation for personal development use; two purposes, perhaps, but not necessarily seen in opposition to one another.

It could be suggested also that both the external and internal environments in which OP operates demand, in general, more ‘compliance’ from staff with institutional requirements. In comparison, the long-standing concept of ‘academic freedom’ at universities creates an environment where institutional requirements are acknowledged at a higher level, but are not always part of common practice or well-accepted across all areas of the institution. The decentralised organisational structures of universities and the large scale numbers of the staff and students as opposed to the centralised structure of OP and smaller staff and student body may mean that decisions from OP senior management are less filtered, more direct, and as a result, more likely to be acted and reported on.

Nationally, universities have greater autonomy than polytechnics around funding models and monitoring regimes. They are less reliant on EFTS funding due to a wider variety of other funding sources from research grants and international students. These variations create unique cultures where different expectations are placed upon staff in order for institutions to be able to fulfil requirements and commitments to government. This may account for the differences in perceptions between OP staff and the staff of the two universities that were found in the data collected for this project.

4.3.3.5 Evaluation/appraisal instruments and systems

Perceptions about the quality of evaluation/appraisal instruments and systems were drawn mainly from the questionnaire comments and interviews, as none of the Likert-scale questions in the questionnaire asked about these views directly. The common themes shown below arise from the perception questions in Section B of the questionnaire, that is, Q9 to Q22. The full set of comment data for Section B of the questionnaire, their theme and sub-theme coding, theme coding by institution and examples of comments categorised within themes and sub-themes can be found in Appendix 6 to Appendix 13. Presentation of the comment data (section 4.3.3.5.1) is followed by an overview of the data drawn from the interviews (section 4.3.3.5.2).

4.3.3.5.1 Evaluation processes and instruments and institutional uses of evaluation – comments from the questionnaire

Outlined below, direct responses to Section B of the questionnaire expressed criticisms of the quality of the feedback in relation to (a) the evaluation instruments and process; and (b) the source of the data, the students. Less direct are the common themes in (c) showing a positive perception of the quality of the data. These showed a positive perception of the data but the perception of quality can only be implied.

The comments from Section B were extensive and they are summarised below in note form. The percentages specified below show the percentage range of the themes across all the Section B questions that contained these types of comments. Each common criticism outlined below is part of a broader theme and so the percentage figure does not represent the percentage of staff that made that particular comment but that the comment fell within a group of related comments. This means that the number of respondents who made these specific comments was not calculated, as the theme coding did not separate them to that level of detail. Rather, of all the comments made, *no more than* the higher percentage figures in each range were these kinds of comments. This gives an indication that the issues listed below are not of concern to a large majority of staff. However, it is useful to know the criticisms as they do represent real concerns of the group of staff who made them. Full analysis of Section B responses appears in Appendix 6 to Appendix 13.

(a) The following were the common criticisms of the evaluation instruments and process. Between one per cent and 14 per cent of the comments covered these themes (combined institutions):

- questionnaire design faults
- questionnaires are too broad/blunt an instrument
- need more choice/flexibility with questions
- limited information available from quantitative data as opposed to the more useful qualitative comments.

Concerns were expressed about statistical methods used and process weaknesses. Between one per cent and 15 per cent of the comments covered these themes (combined institutions).

The comments included the following issues, which were highlighted by respondents contributing to why they considered statistical methods and processes surrounding student evaluation as weaknesses:

- validity, reliability, self-selection bias, non-response bias
- poor response rates
- not an appropriate/valid method for gathering data for some teaching situations (*e.g.* small classes, one-to-one teaching, clinical, supervision)
- system vulnerable to manipulation/error
- results can be affected by timing/frequency of questionnaires
- medium used can affect results (*e.g.* paper and online surveys).

Others commented on concerns with interpretation of student evaluation data. Between one per cent and 14 per cent of the comments covered these themes (combined institutions).

These comments included:

- data gathered is difficult to interpret
- need comparative/contextual data
- data is not used appropriately (weighted too highly, context not considered)
- not suitable to be used for judgement or as a guide to teaching quality.

Other evaluation methods were seen as important too, and that the institution's centralised student evaluation system was not enough to provide a good picture of teaching and courses. Between two per cent and nine per cent of the comments covered these themes (combined institutions):

- should be used in conjunction with, or prefer to use, other evaluation methods.

(b) The second group of comments covered common criticisms of the source of the data, namely, the students.

Students were seen by some as being unable to judge teaching. Between two per cent and 15 per cent of the comments covered these themes (combined institutions):

- students are not qualified/able to judge course content
- not teaching experts; lack experience
- too soon to appreciate the benefits of the teaching or recognise key skills
- not yet capable of valuing their learning experience
- not aware of the bigger picture
- need to balance what students want with what they need (as assessed by teacher/discipline)
- students judge a teacher on factors outside the teachers' control.

Poor quality of feedback was another criticism. Between two per cent and 15 per cent of the comments made covered these themes (combined institutions):

- feedback not practical/possible/constructive
- responses are contradictory/not serious/do not answer the important questions
- over-evaluation reduces quality of responses.

Other comments pointed out that students are too influenced by other factors, and therefore their feedback was not valued. Between two per cent and 15 per cent of the comments covered these themes (combined institutions):

- prone to bias
- results can be skewed by emotional/retaliatory/disgruntled responses
- students base their responses on enjoyment/popularity/entertainment, not quality of teaching/course
- react to difficulty level of course
- resistant to new or challenging teaching techniques
- students have a negative attitude to compulsory papers therefore will comment negatively
- lack commitment to their education (want to be spoon-fed)

- differences between student cohorts (e.g. cultural, ethnic) therefore feedback from one less applicable to another
- not all students learn the same way.

(c) The third group of common themes were those that showed a positive perception of the quality of the data. Because the indications of positive perceptions are less direct, care should be taken drawing too many conclusions. However, it is worth noting some aspects that show positive views of the process and the feedback data provided by student evaluations.

In Q21 of the questionnaire (see Appendix 12) respondents were asked to comment on what they would recommend if they were able to decide the future of student evaluation/appraisal. Fourteen per cent of the comments (166) would keep the status quo. They were positive about the current system in their respective institutions and found the data important/helpful/essential. One implication is that they consider the quality of the data adequate for their use. This was also reflected in Q22 (see Appendix 13), which asked for general comments about student evaluations/appraisals. Four per cent of the comments were on the administrative strengths of the system and eight per cent considered the process to be good/valuable. Below are some examples of comments from these groups in Q21 and Q22.

To keep it definitely as it does help contribute to the idea that students are our best critics and commentators about our teaching, which they are in my opinion. They are objective (they aren't colleagues trying to impress); they are always honest; and often they clearly want to help improve you so that other students benefit. (Q21, sub-theme 1a, OU, senior teaching position, 16–20 years' tertiary teaching experience, permanent position, Humanities.)

Overall, an important mechanism I feel that is underutilized by the university! (Q22, sub-theme 1b, WU, senior teaching position, 11–15 years tertiary teaching experience, continuing position, Commerce.)

They should continue. They are a useful tool and they help empower the students. (Q21, sub-theme 1a, OP, lecturer teaching position, 0–5 years' tertiary teaching experience, fixed term position, Sciences.)

The ratings question Q19 (see Table 4:20, below) may also indirectly shed some light on staff views of the quality of the data in that it asked *Q19 How effective is your institution's centralised evaluation/appraisal system in gathering useful/meaningful student data for you?*

Table 4:20 Combined Total Ratings for Q19

	1&2 %	middle ground %	4&5 %
OU	59%	25%	16%
WU	63%	21%	15%
OP	53%	31%	16%
All	60%	25%	16%

Table 4:20 shows that over half the respondents indicated that they found the centralised system effective at gathering useful/meaningful data for them (1 or 2 rating), while 16 per cent on average found the centralised system not effective (4 or 5 rating). Again, this is not a direct comment on the quality of the data gathered through centralised student evaluation systems, but does imply that the data gathered are considered good enough to be useful/meaningful to staff. The following quotes illustrate some respondents' positive views of the quality of the data. These are taken from Q20 (see Appendix 11), which asked respondents for comments about their responses to Q19.

It gathers the sorts of information I want and have sought in the past, with the advantage that a lot of the work I had to do myself is now done for me. (Q20, sub-theme 1b, OU, senior teaching position, 21+ years' tertiary teaching experience, Health Sciences.)

Have consistently been given constructive feedback from the survey responses. (Q20, sub-theme 1a, WU, lecturer teaching position, 6–10 years' tertiary teaching experience, continuing position, Sciences.)

A bit slow, but once it arrives, good. (Q20, sub-theme 1b, OP, senior teaching position, 16–20 years' tertiary teaching experience, permanent position, Humanities.)

The recent research indicated a relatively positive disposition towards evaluation (*e.g.*, Beran & Rokosh, 2009), and accordingly, the criticisms outlined above do not represent a large proportion of the comments (usually between one per cent and 15 per cent). So while these concerns do need to be addressed, they should not be considered widespread issues for all staff. Even so, although the numbers are not high, they are concerns for the staff who made

the comments. To maximise the benefits of the student evaluation/appraisal process for as many staff as possible, these issues need to be carefully considered and addressed by the respective institutions.

Another aspect that was observed during the analysis of the questionnaire comment data was that some teachers expressed the view that the data was of sufficient quality for formative purposes, but not of sufficient quality to be used in a judgemental or summative way, for example, for promotion or other summative purposes. This tension between summative and formative purposes is a common theme throughout the questionnaire comment data.

I think that students' evaluations are a useful tool, bearing in mind that their primary function should be about improving the quality of teaching and learning. I'm not so sure about their function as quality assurance tools. It seems that peer review of entire programmes and external assessment is a better way to achieve quality assurance. (Q22, sub-theme 2f, OU, lecturer teaching position, 11–15 years' tertiary teaching experience, permanent position, Health Sciences.)

I would continue an appraisal, probably increase the amount of open questions, and review whether data should be used for promotion. There is considerable tension between this and gathering data for the enhancement of teaching. (Q21, sub-theme 4a, WU, lecturer teaching position, 16–20 years' tertiary teaching experience, continuing position, Humanities.)

4.3.3.5.2 Evaluation processes and instruments and institutional uses of evaluation – comments from the interviews

4.3.3.5.2.1 University of Otago

A strong unease expressed in the OU interviews concerned student limitations. The most repeated view was the potential unreliability of student evaluation feedback, because students are seen to put too high a premium on perceived irrelevancies such as likeability, entertainment and popularity. Nine respondents indicated, to varying degrees, that these personality traits of academics may distort student judgement. By implication, the personal and relational dimensions of teaching are separated from substance and content dimensions in the views expressed here. Examples of comments include:

I know many of them would know if they found you likeable. It's a popularity contest, a lot of it.

It's a popularity contest. I'm embarrassed to be using them.

I guess there's a risk that students will overrate entertainment value and underrate sort of solid factual content.

You should treat them with a pinch of salt. Students don't have a long-term perspective. Personality is a big factor.

Other factors were also commented on as distorting students' judgements about teaching and learning. These include a generalised sense that a range of factors that prevent students from seeing the whole picture. Limiting factors noted were a lack of maturity, an inability to grasp the course learning goals, the difficulty of a paper and not having the advantage of perspective or retrospective insight.

Some OU interviewees also linked perceptions about the unreliability of student evaluations to the belief that the evaluations can be, and often are, frequently manipulated by academics to promote their career advancement. Eight OU interviewees articulated views in this vein. Examples of these comments include:

Some academics choose their best evaluations.

The student evaluations instrument is unreliable, contrived and manipulated.

So I tend to ask questions where I know I'll get a good grade.

The range of questions means it's open to manipulation to make the teacher look good.

Negative evaluations can affect your advancement. And for that reason people are careful to choose questions that are more likely to yield a positive response.

Another recurring objection expressed by OU interviewees related to the perception that there are flaws in the evaluation instrument or the institution's evaluation systems. Eight interviewees expressed reservations in this respect. Comments include:

Evaluations are unreliable for a small postgraduate class.

Statistical concerns...

University forms are a blunt instrument.

Could improve system with less but more focussed questions and more flexibility for people to write their own questions.

People are more careful to choose questions that are more likely to yield a positive response.

Underlying the perception that evaluations are a dubious indicator of good teaching, was the reference by OU interviewees to the use of evaluations/appraisals by the organisation for quality and career advancement. These comments, which tended to focus on manipulation suggest, for the interviewees making them, that the use of student evaluations to demonstrate teaching competence to the organisation undermines their usefulness for professional development. The negative impact of the institutional uses of evaluation for promotion was explicitly mentioned by eight OU interview participants. Examples of comments include:

The institution tries to do too much with this limited data.

So I tend to use the questions where I know I'll get good grades. Especially when using for promotion.

The summative use undermines the formative.

So we know full well that the results of teaching evaluation could be used against us.

Corresponding to reservations about the formal student evaluation system, a number of interviewees expressed the view that formal evaluations need to be complemented by other evaluation strategies. This view was stated by six interviewees and, as will be seen at OP, suggests that the use of multiple forms of evaluation may enable teachers to see formal evaluation as part of a total picture, or, in some instances, disregard the formal system entirely.

Finally, as participants from the other institutions mentioned in their interviews, the timing of the evaluations was also raised as a problem (three).

4.3.3.5.2.2 The University of Waikato

Eleven WU respondents linked their reservations about student evaluation/appraisal to what they saw as shortcomings in student approaches to, and practices of, evaluation. Unlike their

interview counterparts at OU, WU interviewees did not point to one particular reason why students' judgements may be untrustworthy, but instead, demonstrated a mix of individual beliefs and sentiments about the shortcomings of students' views. Only one interviewee referred to students' tendency to confuse likeability or popularity with educational value, as opposed to nine at OU who held this view. Perceptions about the limitations of students' judgement included: the view that students cannot see the whole picture (two); the idea that students' views are influenced by the payment of tuition fees (two); the observation that many students do not attend class (two); the belief that a difficult paper skews students' evaluations (one); the notion that students use the evaluations to voice personal grievances with the lecturer (one); that students' comments were superficial and vacuous (one); and perceptions that students were constrained in responses because of cultural factors (one). Only one interviewee at WU expressed the concern that teachers could manipulate the evaluation instrument, saying:

The fact that you're allowed to choose your own questions, means that you can direct students to answer questions about your strengths, rather than your weaknesses.

Interviewees at WU raised concerns about flaws in the evaluation instrument or the institution's student evaluation systems. Eight WU educators raised questions about the instrument and its validity (eight concerns of this nature were also articulated by OU interviewees). Issues raised included limitations of the questions, size of the sample completing the evaluations and the need for more open-ended questions. A number of interviewees indicated that they used other forms of evaluation. In some instances they saw personal feedback as a good complement to the formal evaluation system, while for others, personal feedback was seen as a way to strengthen views about the limits of formal student evaluations.

The institutional use of evaluations did not feature as strongly as a negative in WU perceptions of evaluations as it did in the OU findings. Three WU interviewees voiced concerns about the use of evaluations for promotion, as opposed to eight at OU. However, three other WU interviewees were concerned that the student evaluations were primarily used by the institution as a promotion tool and believed that the university needed systems in place that would encourage and enable academics to utilise evaluations to improve their teaching.

An associated concern at WU was the timing of the student evaluations, which was identified by nine interviewees as a barrier to effective use of and benefits from the evaluation system.

4.3.3.5.2.3 Otago Polytechnic

Concerns about evaluation processes expressed by OP interviewees included the perception that students are unduly influenced by popularity and likeability as opposed to substantial learning. Personal and emotional factors were strongly articulated by some teachers.

Although these kinds of comments were only offered by six people, they were presented with some intensity. Other familiar reservations were expressed by some interviewees, including issues with the evaluations instrument and the idea that evaluations are open to manipulation by staff members.

Timing of evaluations was raised as a flaw in the evaluation system by five OP interviewees. They suggested that the timing meant that the feedback was often received too late to be used for the current cohort of students.

A theme that recurred in the comments of a number of respondents (five), was different attitudes to, and use of, course evaluations as opposed to teacher evaluations. A repeated notion was a readiness to discuss course feedback more widely with colleagues and students. Teacher evaluations, on the other hand, were perceived as more private.

Contextual factors also influenced the way OP educators perceived the usefulness of the centralised formal student evaluations system. One strong theme that emerged from the OP interviews in this section was that small classes and the many practical learning sessions meant that the tutors were in regular contact with their students and were observing them and finding out about their learning in an on-going way (five). Comments about the relevance of this for attitudes to evaluation varied. In some instances, it meant that respondents were more relaxed about student evaluations, that there were few surprises and that evaluation tended to be a confirmation or a validation of what they already knew. In some instances, this led respondents to pay less attention to the formal evaluations. Examples of comments include:

If you interact with your students it should be a confirmation of what you know already.

The best feedback is from them directly face to face. So I guess I rely on that because I do think I get more honest feedback face to face.

I think that you form a different relationship with them...I think they are more forthright and open to discussion.

Associated with the regular contact linked to small classes and practical sessions was the reference to using a number of methods of formative evaluation by nine interviewees. The effect on attitude to formal evaluations was along similar lines to that discussed in relation to contact in small classes.

Concerns about the negative effects of the institutional uses of student evaluation did not surface as a strong theme. Only two people commented on this issue; one mentioning that it could be a concern. The other comment in relation to the institutional use of evaluations was the view that the full developmental use of evaluations was not realised, because the institution was only focussed on the quality monitoring aspects:

Often I think it's a check box afterthought by the organisation.

Some interviewees noted that the different students, in different courses and levels, meant that there was a wide variation in terms of the kind of evaluation students were able to make.

Unlike in the universities, discipline research requirements were not mentioned by any OP interviewees as barriers to teaching or as a factor undermining the relevance of student evaluations.

4.3.3.5.2.4 Overview of the interview data highlighting differences between Otago Polytechnic and the two universities

The analysis of the interview data showed that contextual factors result in OP standing out from the two universities in a number of ways. The following items highlight those differences. The number of interview responses in which the differences were mentioned is included in parentheses.

- the focus of much teaching is practical and workplace orientated (11)
- close attention to the students' comments about educational role and on what they expect their students to be able to learn/do; very high number of interviewees focussed on outcomes for students in their educator role (12)

- very student-referenced in their discussions (12 directly)
- evidence of use of other forms of evaluation (nine)
- small class sizes and therefore close connection with students means that teachers know about their students' learning progress in an on-going way (five)
- about a third of the teachers talk about the challenges of moving from a practitioner to an educator role and trying to get this right (seven)
- no comment on the impact of discipline research on teaching, its connection with students and issues of the institution valuing research over teaching (zero)
- discussion of diverse student body and impact on teaching and evaluations (three)
- not much emphasis on problem of institutional use of evaluations (two)
- clear separation of course and teacher evaluations and their use (five)
- stronger evidence of engagement with responses and using feedback to inform teaching (17)
- some evidence of reporting on feedback and responses to student, although mainly on the course evaluations, rather than teacher evaluations (five)
- greater evidence of recognition of a higher education language (seven).

The following list is of the features of commonality that the interviews highlighted between the OP and the universities:

- the perceived problem of likeability/popularity versus genuine learning benefits (five)
- sensitivity to personal or negative comments or the one negative comment syndrome (six)
- problems regarding the timing of evaluations both in terms of using them to inform current teaching and giving feedback to students on their comments
- institutional messages about evaluations, though less than in the universities' responses
- questions/reservations about the validity of the instrument and the usefulness of the questions
- possible manipulation of evaluations by teachers
- mixed responses about student competency to judge (does not necessarily match with foregrounding of students in OP thinking).

4.3.4 Conclusions

In summary, the perceptions of the quality of the data expressed by participants were widely varied, although the size of the group of respondents who viewed evaluation/appraisal tools, processes and systems negatively is smaller than was expected. Many of the concerns and criticisms expressed by the participants in the interviews and through the comment data in Section B of the questionnaire have arisen in the literature and do echo some of the common ‘myths’ around student evaluation often encountered by teaching staff. The evidence indicating a positive perception of the quality of student evaluation data does relate to a much larger group of participants, but any claim made needs to be treated with caution, as these claims are only implicit.

4.3.5 Question 3: How do tertiary teachers engage with evaluation results and students' feedback?

As discussed in the literature review (section 2.1.3.3) there seems to be a connection between teachers’ beliefs around their teaching and how they engage with evaluations/appraisals. These beliefs influence their relationship with their students and their engagement with the evaluation/appraisal process, but also the institutions’ use of evaluation/appraisal can influence, reinforce and provide barriers to their views of this process. Research question 3 *how do tertiary teachers engage with evaluation results and students’ feedback?* sought to investigate these notions.

In order to discuss findings in answer to research question 3, a series of sub-questions are used to frame the presentation of evidence drawn from the questionnaire. These four sub-questions are:

- a) What do teachers think of the feedback they get from students through the formal evaluation/appraisal system?
- b) How do lecturers interpret the data?
- c) How do teachers feed the outcomes back to students?
- d) How do lecturers use their evaluation/appraisal for professional development?

The sub-questions, discussed in section 4.3.5.1, model effective use of evaluation/appraisal to inform teaching practice and student learning. In addition, data from the interviews are used

to provide further illustration and exploration of academics' engagement with evaluations. Interview data are discussed in section 4.3.5.2.

4.3.5.1 Teacher engagement with evaluation – evidence from the questionnaire

- a) What do teachers think of the feedback they get from students through the formal evaluation/appraisal system?

A first step is to look at whether or not teachers think it is worthwhile collecting evaluation/appraisal data. As was discussed in section 4.3.1, this is elucidated by the summative responses to Q17 (*Do you consider it worthwhile to gather evaluation/appraisal data about teaching and courses/papers?*) repeated here in Table 4:21 below.

Table 4:21 Responses to Q17 (*Do you consider it personally worthwhile to gather evaluation/appraisal data about teaching and courses/papers?*)

Institution	Number of responses	Very worthwhile (1 & 2)	Middle ground (3)	Not at all worthwhile (4 & 5)
OU	n=567, 85%	70%	18%	12%
WU	n=237, 98%	74%	15%	11%
OP	n=139, 91%	85%	10%	5%
All	n=943, 89%	73%	16%	11%

As Table 4:21 shows, out of 943 responses to this question, 73 per cent from the three institutions responded that this was a very worthwhile endeavour (rating 1 and 2) with OP having the highest proportion of respondents at 85 per cent. Both the universities were lower (OU, 70 per cent and WU, 74 per cent).

The thematic analysis of the comments (Q18) made by those teachers who thought it worthwhile to collect evaluation/appraisal data (447 responses – 56 per cent) identified two main sub-themes – first, to provide information to inform paper and teacher development (19 per cent of responses) and second, to help identify student learning needs/experiences (also 19 per cent of responses). This second theme mentioned *giving students a voice* and *part of the relationship with students* as a strong reason for collecting this information. The larger number of comments that fitted into this theme came from WU (23 per cent), while both OP and OU were fairly consistent at 19 per cent and 18 per cent respectively. The distribution of the Q18 comment data around the two themes can be found in Appendix 10.

Q2 (*Please identify why you use student evaluations/appraisals*) of the questionnaire was discussed in section 4.3.1 in terms of overall responses. Q2 data is presented again as a way of examining variations in responses on an institutional basis and with a view to addressing research question 3 about teachers' engagement with student evaluations/appraisals specifically.

As the majority of respondents to Q17 believed that the collection of evaluation/appraisal data is worthwhile, it is how teachers use the data that is revealing. Question 2 asked teachers why they use evaluations/appraisals and to respond by selecting 'yes' or 'no' against a list of possible reasons. Table 4:22 below lists the possible uses, as outlined in the questionnaire.

How each institution ranked each possible use, from 1–8 is also included.

Table 4:22 Summary of 'Yes' Responses to Q2 with Proportion of Responses (Rankings 1–8) Made by Teachers at Each Institution

Why use evaluations/appraisals?	OU	WU	OP	No. of total respondents
<i>2a Because it is required</i>	4 (n=462, 82%)	3= (n=193, 85%)	1 (n=132, 97%)	n=924, 87%
<i>2b For my professional development</i>	3 (n=499, 88%)	5 (n=179, 83%)	4 (n=115, 88%)	n=915, 86%
<i>2c For my promotion</i>	5 (n=459, 81%)	3= (n=191, 85%)	7 (n=70, 65%)	n=898, 84%
<i>2d For professional goal setting/salary review/confirmation</i>	6 (n=338, 67%)	6 (n=130, 64%)	6 (n=76, 70%)	n=820, 77%
<i>2e For paper/course refinements</i>	2 (n=507, 89%)	2 (n=191, 87%)	3 (n=128, 94%)	n=925, 87%
<i>2f To get feedback on students' learning experiences</i>	1 (n=543, 92%)	1 (n=216, 95%)	2 (n=135, 96%)	n=959, 90%
<i>2g To provide feedback to my students</i>	7 (n= 190, 41%)	8 (n=86, 43%)	8 (n=65, 63%)	n=767, 72%
<i>2h To report on quality matters</i>	8 (n=172, 38%)	7 (n=92, 47%)	5 (n=81, 74%)	n=757, 71%

Of the eight suggested uses, Q2a to Q2h, a ranking of 1 indicates which of Q2a to Q2h was selected most often by respondents from a particular institution, while a ranking of 8 indicates which of Q2a to Q2h was selected least often from the institution. The 'rankings' were obtained from the proportion of responses made by teachers at each institution.

For this question, teachers at OU (92 per cent) and WU (95 per cent) indicated that they primarily used evaluations/appraisals (Q2f) *to get feedback on my students' learning experience*. Meanwhile, teachers from OP ranked this in second place, with 96 per cent of teachers using evaluations/appraisals for this purpose. The use of evaluations/appraisals ranked highest (but only just) by OP respondents was Q2a, *because it is required*, with almost 97 per cent of the teachers at OP responding 'yes' to this suggested use. This may be a reflection of the clear expectations at OP around collecting evaluation/appraisal data (see Table 2:5). Interesting to note, however, are the relative scores of OU and WU in comparison with OP. The number of teachers who responded 'yes' to 2a at WU also made this the joint third most common use of evaluation/appraisal data, whilst OU teachers did not rate this use quite as highly (ranked fourth). This may be an indication of how evaluation/appraisal data are used by the two universities. A good example is how the two universities use evaluation/appraisal data for promotion purposes. OU requires the submission of a teaching portfolio for promotion (see section 2.3.1.1.4), which includes a range of data including student evaluations data. At WU, the only evidence of teaching required for the promotions process is the appraisal data (see section 2.3.1.2.4). For WU, responses to Q2c *for my promotion* is ranked third along with Q2a *because it is required*, also ranked third. This is possibly an indicator of the reliance placed on appraisal as the sole mechanism for the judging of teaching in promotion applications at WU.

While the majority of teachers who responded to Q17 think it is personally worthwhile to collect evaluation/appraisal data, perhaps as a general principle, from the thematic analysis of the comments made in answer to Q18 as discussed in section 4.3.1, there is a definite group of teachers who do not share this view (Q18, Theme 2, n=313, 39 per cent) in reference to issues and concerns, which tend to limit their personal sense of worth of student evaluation/appraisals. There are two predominant sub-themes – 2c because the *limitations of the current system* reduce its personal worth (Q18, sub-theme 2c, n=115, 14 per cent), for example:

It would be more valuable if we had a better system for doing it e.g., you could design your own questions; and questions were more suited to wider variety of teaching contexts such as health service staff. (Q18, sub-theme 2c, OU, senior teaching

position, 16–20 years' tertiary teaching experience, permanent position, Health Sciences.)

The constant requirement to do it for every paper every time it is taught is tedious and annoying for staff but I think even more so for a lot of students. (Q18, sub-theme 2c, WU, lecturer teaching position, 21+ years' tertiary teaching experience, continuing position, Commerce.)

I feel it is very worthwhile to get feedback from students but the surveys don't facilitate good usable data in my experience. Over the course of five years I have had survey results from hundreds of students and it is usually all the same. (OP, lecturer teaching position, 6–10 years' tertiary teaching experience, permanent, Sciences.)

and Q2d because the *quality of the students responses are questionable* (Q18, sub-theme 2d, n=83, 10 per cent). Comments categorised under sub-theme 2d mention *frequency; and students can't judge teacher effectiveness*. For example:

Mostly the information is too difficult to interpret in any coherent way; too contradictory; less a comment on my teaching than on their feelings. I have found very little on evaluation forms that has given me clear feedback about things it's in my power to change (or continue). There are more effective ways of evaluating teaching. The student forms evaluate their 'experience'. (Q18, sub-theme 2d, OU, lecturer teaching position, 16–20 years' tertiary teaching experience, permanent position, Humanities.)

I have a number of examples of receiving meaningless feedback from students, e.g. when questions have been asked by coordinators about all members of a team and I have received feedback from students on tasks I was not part of. This does make me question how reliable the results are on all appraisals. (Q18, sub-theme 2d, WU, lecturer teaching position, 11–15 years' tertiary teaching experience, continuing position, Humanities.)

A lot depends on the type of course I teach. It is not popular with many students but is compulsory for their degree therefore the type of feedback is often negative purely due to the students disliking the subject and not because it hasn't been taught well.

(Q18, sub-theme 2d, OP, lecturer teaching position, 6–10 years' tertiary teaching experience, permanent position, Health Sciences.)

Doubts about the capacity of students to be able to judge teaching were also expressed in the interviews. Thirteen of the 20 interviewees at OP indicated some ambivalence about students' evaluative competencies, while at OU, the largest group of respondents (12) fell into a middle group who expressed views ranging from some reservations to serious scepticism.

Additionally, three interviewees were emphatic in their views that students could not evaluate their learning experiences. This means that 15 out of 20 interviewees at OU questioned students' evaluation competence. At the WU, ten interviewees expressed support for the idea that students could make judgements about their learning experiences ranging from emphatically positive to very positive with slight reservations. The other ten included a middle group with varying degrees of reservation (seven), to three who were strongly negative about students' capacity to judge their learning. Thus over the three institutions, a total of 32 interviewees expressed ambivalence in varying degrees about students' capacity to judge the teaching and their learning experiences. A further six interviewees were decisively negative about students' competence in these respects.

What is evident here is that these teachers perceive that student feedback is limited (and cite reasons limiting their sense of personal worth of student evaluation/appraisal), but also that they, as teachers, do not see the place of engaging with the *students* around evaluation/appraisal. Rather, engagement, if it happens, whether negative or positive, limiting or enhancing, is with themselves as staff and with the tools, instruments and processes surrounding the student evaluation/appraisal system at their institution. This could indicate a lack of ownership of the evaluation/appraisal process and a perception that student evaluation/appraisal is something that is 'done' to teachers by the institution, rather than seeing evaluation/appraisal as integral part of their teaching practice that involves students.

A slightly different perspective comes from the thematic analysis of comments for Q21 *if you were able to decide the future of student evaluation/appraisal at your institution, what would be your decision and why?* The themes that were identified and outlined in detail in Appendix 12 include keeping the status quo (Theme 1, 30 per cent), taking a developmental focus (Theme 2, 32 per cent), suggestions for administrative changes (Theme 3, 11 per cent) and

comments on institutional use (Theme 4, 23 per cent). The percentage contribution of thematic comments is broken down into each institution in Appendix 12.

Question 21 sub-theme 2e is made up of comments about the role of the student (n=61, five per cent) and possible measures that could be taken to guide and educate students through the evaluation/appraisal process. Comments under this sub-theme were broadly about: *educate students about value and consequences, better buy-in from students needed; need to help the students provide more useful information; bias/popularity effects; and close the feedback loop so students treat it more seriously*. Question 21 comments, along with comments from other parts of Section B of the questionnaire, were included in the more wide-ranging discussion in section 4.3.3.5.

b) How do lecturers interpret the data?

While the data shown in Table 4:22 show that most teachers think it is a worthwhile endeavour to collect evaluation/appraisal data, and that they use it predominantly to obtain feedback on their students' learning experiences (in answer to Q2f), further insights into how teachers engage with evaluation/appraisal can be gained by considering what the teachers actually do with the evaluation/appraisal results once they have received them. Question 4 in the questionnaire specifically asked this (*Q4 when you receive the results from your student evaluations/appraisals, do you...*). The aggregate data, which was already presented in section 4.3.1, is once again presented in Table 4:23 below, but with the addition of overall rankings of importance, as determined by the number of responses to each Q4 sub-question. Rankings are from 1–7 to match the seven sub-questions 4a to 4g, with ranking 1 indicating the sub-question that received the highest percentage of 1 and 2 responses and ranking 7 indicating the sub-question which received the least.

Table 4:23 Summary of Overall Ranked Responses to Q4a-4g

Q4a Sub-questions	Ranking	Institution	1=Always (% 1&2)	Middle ground (% 3)	5=Never (% 4&5)
<i>4b Read the open question comments/ responses made by the students?</i>	1	OU (n=622)	96%	2%	2%
	1	WU (n=237)	91%	4%	5%
	1	OP (n=135)	96%	2%	2%
	1	ALL (n=994)	95%	2%	3%
<i>4a Spend time going over the data and responses?</i>	2	OU (n=621)	89%	7%	4%
	2	WU (n=235)	81%	14%	5%
	2	OP (n=134)	91%	5%	4%
	2	ALL (n=990)	87%	8%	4%
<i>4g Actively look for feedback about teaching and assessment?</i>	4	OU (n=610)	75%	13%	12%
	3	WU (n=236)	76%	14%	10%
	3	OP (n=133)	84%	9%	7%
	3=	ALL (n=797)	77%	13%	11%
<i>4d Compare the data with previous evaluations/ appraisals?</i>	3	OU (n=614)	81%	11%	7%
	4	WU (n=231)	74%	13%	13%
	4	OP (n=128)	65%	21%	14%
	3=	ALL (973)	77%	13%	10%
<i>4e Discuss the results with colleagues/ teaching team?</i>	5	OU (n=615)	48%	24%	27%
	5	WU (n=231)	36%	30%	33%
	5	OP (n=98)	64%	22%	13%
	5	ALL (n=944)	47%	25%	27%
<i>4c Provide students with feedback on the results?</i>	6	OU (n=603)	14%	14%	71%
	6	WU (n=230)	16%	15%	70%
	6	OP (n=130)	26%	19%	55%
	6	ALL (n=963)	16%	15%	69%

<i>4f Seek assistance with interpreting the results from others?</i>	7	OU (n=602)	12%	15%	73%
	7	WU (n=226)	8%	10%	82%
	7	OP (n=92)	25%	20%	55%
	7	ALL (n=920)	12%	14%	74%

Table 4:23 also shows that the lowest ranking, 7, is given to Q4f *seeking assistance with interpreting the results from others*. This issue of seeking help with interpretation has been identified in responses to a number of questions in the questionnaire. For example, one of the themes that was identified in the accompanying Q18 comment data (see Appendix 10) was difficulties in relation to interpretation and effective use of data (Q18 sub-theme 2f, n=14, 2 per cent). Within the Q18 sub-theme 2f, a number of comments related to the fact that *interpretation of feedback was difficult, especially criticisms*. The question of interpretation of student evaluation/appraisal data is one that seems, from these comments, to be recognised as a problem by a small group of teachers, and small numbers of teachers say they seek assistance with interpretation (see 4f in Table 4:23). For example, the analysis of comments in answer to Q10 (*Please explain your answer to Q9*) identified that some teachers said that evaluation/appraisal data did not influence their teaching decisions because the quality of the data was not good enough (Q10, sub-theme 2a, n=107, 13 per cent – see Appendix 6). For example,

It is my perception that the evaluations are largely a measure of popularity rather than of learning/teaching. I believe that the students are generally not in a position, at the time the evaluation is set, to comment on their learning outcomes or my teaching. (Q10, sub-theme 2a, OU, lecturer teaching position, 16–20 years' tertiary teaching experience, confirmation position, Health Sciences.)

First, it depends on the level assessed for the teaching appraisals. For instance, most first-year students are not discerning. They don't really have a clue about assessing teaching methods or, indeed, about content...Second, the appraisals themselves are flawed: one might get a good appraisal simply because the workload of a paper is low, not necessarily on whether the students get a comprehensive overview of a subject (how would they know anyway?) or whether they have learnt a set of key skills. (Q10, sub-theme 2a, WU, lecturer teaching position, 0–5 years' tertiary teaching experience, continuing position, Humanities.)

Often students' comments are not realistic and unable to be implemented. (Q10, sub-theme 2a, OP, lecturer teaching position, 11–15 years' tertiary teaching experience, permanent position, Health Sciences.)

This theme was particularly prevalent in the comments made by WU respondents with 16 per cent (n=37) of the WU comments being critical about the quality of the data.

A number of the questionnaire responses made suggestions concerning the interpretation of evaluation/appraisal data. For example, from the analysis of comments made in response to Q21 (*if you were able to decide on the future of student evaluation/appraisal at your institution what would be your decision and why?*) a theme emerged around guidance/support for teachers (Theme 2 – Developmental Focus, sub-theme 2c, n=24, two per cent) with comments suggesting *better guidance/support for staff: how to use as a tool for improvements (i.e. educate teachers in formative uses), how to interpret and act on the results* (see Appendix 12). Responses to Q22 (*If you have any general comments to make about your institution's evaluation/appraisal process, please make them here*), also highlighted that an area where improvement was needed was in guidance on interpretation of results (Q22 sub-theme 2d, n=24, five per cent), with some of the comments identifying the *need for benchmarking/comparative data, caution with interpreting the data and realistic expectations, particularly with challenging papers, the importance of teaching context* (see Appendix 13).

c) How do teachers feed the outcomes back to students?

One of the lowest ranked uses of evaluation/appraisal data has been to provide feedback to students (see Q2g, ranking 7 or 8 in Table 4:22 and Q4f, ranking last at 7 in Table 4:23). For Q2g, the two universities had 41–43 per cent of staff indicating that they did provide feedback to students as a reason why they use evaluations/appraisals, with OP much higher at 63 per cent. For Q4f, the two universities had only 14–16 per cent of staff who indicated that they routinely provided feedback to students, with OP being marginally higher at 26 per cent. Further exploration of teachers' perceptions about providing feedback to students can be made by examining the responses to the two-tiered Q6 paired with Q7. The first part, Q6, asked teachers if they communicated with students about feedback, using a 'yes'/'no' scale (*Q6 do you ever communicate with students about their feedback from student*

evaluations/appraisals?). The second part, Q7, was for those who answered ‘yes’ to Q6. Question 7 asked *do you show your students you have taken account of their feedback from the evaluations/appraisals through...* and followed with a series of four prompts, which reflect common mechanisms, identified in the literature review that communication of evaluation/appraisal feedback back to students occurs. They could add any other ways they communicated with students in Q8, which was an open-ended question. The responses to Q6 and Q7 are shown in Table 4:24 below. The first point to notice from the data presented in Table 4:24 is that the numbers of Q7 respondents from both OU and WU are generally greater than those who answered ‘yes’ to Q6, with OP having the higher proportion answering ‘yes’ (55 per cent). It seems that some respondents who answered ‘no’ to Q6, did not move on to Q9 as asked in the instructions, but answered Q7 as well.

In terms of communication by informal discussion with students (Q7b), it is evident from Table 4:24 that OP staff say they engage with this process to a greater extent than staff at either of the universities (75 per cent 1 and 2 from OP, as opposed to 58 per cent from OU and 46 per cent from WU).

Table 4:24 Frequencies of ‘Yes’ Responses to Q6 and Q7

Q6	OU	WU	OP	All
‘yes’ responses	n=231, 37%	n= 81, 34%	n=76, 55%	n=388, 39%

Q7 sub-questions				
	1=Always (% 1&2)	Middle ground (%3)	5=Never (% 4&5)	Count (%) total respondents
<i>7a The paper/course outline</i>				
OU	106 (41%)	39 (15%)	111 (43%)	256 (38%)
WU	57 (40%)	21 (15%)	64 (45%)	142 (59%)
OP	32 (50%)	11 (17%)	21 (33%)	64 (42%)
ALL	195 (42%)	71 (15%)	196 (42%)	462 (43%)
<i>7b Informal discussion with students</i>				
OU	158 (58%)	63 (23%)	52 (19%)	273 (41%)
WU	64 (45%)	33 (23%)	44 (31%)	141 (58%)
OP	57 (75%)	15 (20%)	4 (5%)	76 (50%)
ALL	401 (79%)	111 (23%)	100 (20%)	490 (46%)
<i>7c Paper/course refinements/improvements</i>				
OU	229 (80%)	35 (12%)	23 (8%)	287 (43%)

WU	106 (72%)	18 (12%)	23 (16%)	147 (61%)
OP	66 (92%)	2 (3%)	4 (6%)	72 (47%)
ALL	401 (79%)	55 (11%)	50 (10%)	506 (48%)
<i>7d Departmental communication channels</i>				
OU	51 (20%)	31 (12%)	169 (67%)	251 (37%)
WU	14 (11%)	14 (11%)	101 (78%)	129 (53%)
OP	17 (28%)	11 (18%)	33 (54%)	61 (40%)
ALL	82 (19%)	56 (13%)	303 (69%)	441 (41%)

While a greater proportion of OP respondents claimed they engage with evaluations/appraisals by feeding back to students in comparison with the other two institutions, respondents from all three institutions appear to be in agreement with the preferred means of feedback. Table 4:24 shows that the most used method for providing feedback, for all institutions, is via the paper/course refinement (Q7c), which is 79 per cent of those who responded to this question. This is similar to the Q2c result shown in Table 4:22 which attracted a high ‘yes’ response overall. The least used mechanism shown in Table 4:24 is 7d (*departmental communication channels*). Only about a fifth of the staff who feed back to students use this mechanism frequently. It is interesting that at WU, only 40 per cent of teachers said that they frequently provided feedback via the paper outline (Q7), even though it is a requirement that all paper outlines contain some feedback of the previous appraisal and any changes that have occurred as a consequence (see section 2.3.1.2.2). The percentage is similar for OU (41 per cent) who have no such requirement. This mechanism is used more by OP staff who, like OU, have no such requirement (50 per cent).

The thematic analysis of comments made in response to Q8 (*if there are other ways you communicate with your students about evaluation/appraisal results, please outline them here*) shows that one of the main factors for not communicating with students about their feedback is that of timing. When most evaluation/appraisal data becomes available to staff, the cohort of students who provided the feedback has moved on. For example:

Once students have completed an evaluation for my course, they move on and are not taught in this course again. I do talk with subsequent students about evaluation responses in general terms, but it is not a direct feedback loop to the same students who did the evaluation. (Q8, OU, lecturer teaching position, 11–15 years’ tertiary teaching experience, permanent position, Health Sciences.)

Usually the feedback is too late for the students currently in your class so there is no 'direct' feedback to them. As they have told me, end-of-semester evaluations are of no benefit to us, they only help students who come after us. This is why I instituted the early/mid semester written 'how are we going?' appraisal because they told me they like this better because it may lead to immediate changes that will help them. (Q8, WU, senior teaching position, 16–20 years' tertiary teaching experience, continuing position, Humanities.)

I tell one group of students what the feedback was from the previous group of students, when I remember. I usually only collect evaluations from students formally at the end of the course and therefore don't provide feedback to the current years' students. (Q8, OP, senior teaching position, 16–20 years' tertiary teaching experience, permanent position, Health Sciences.)

However, the comments of a small group of teachers showed that they use other forms of informal evaluation to gain feedback from, and 'feed forward' to, students on their learning and the paper. There was also a small number of comments made concerning the confidentiality about providing feedback to students. For example:

Most often I administer the centralised questionnaires at the end of semester and so provide no direct feedback to students who are long gone by the time I receive the analysis. I do provide students feedback from my more informal questionnaires. (Q8, OU, lecturer teaching position, 16–20 years' tertiary teaching experience, permanent position, Health Sciences.)

Because of the timing of the appraisals (end-of-paper), it is not normally possible to have discussions with the students who provided the feedback and open-ended comments. There is a risk of confidentiality (anonymity) being compromised. However, these comments may translate into revisions of the next offering of the same paper. (Q8, WU, lecturer teaching position, 21+ years' tertiary teaching experience, continuing position, Humanities.)

I think it's worthwhile noting that formal evaluations are not the only way that I seek and get feedback and that the process of getting and responding to student feedback

is really a never-ending chain. (Q8, OP, senior teaching position, 0–5 years' tertiary teaching experience, permanent position, Humanities.)

This issue of timing also came through in the thematic analyses of Q21 and Q22 (see Appendix 12 and Appendix 13). In response to Q21 (*if you were able to decide on the future of student evaluation/appraisal at your institution, what would be your decision and why?*) there were a number of themes in relation to feeding back to students. Under Theme 2 (Developmental focus), there were comments made concerning the role of students (Q21, sub-theme 2e, n=61, five per cent of respondents) and in particular the importance of *closing the feedback loop so students treat the evaluation/appraisal process more seriously.*

I definitely would not get rid of student evaluation, but I think it could be improved. I don't think that young students are really aware of their power, and haven't been educated in the use of student evaluations. For instance, the first time they ever do an evaluation, they should be informed as to why it is done, what the results mean, what happens to the results. (Q21, sub-theme 2e, OU, lecturer teaching position, 16–20 years' tertiary teaching experience, permanent position, Health Sciences.)

First, I would try to ensure that those taking part in appraisals had some education on the process itself. Then I would institute a two-way process. Students should engage in self-appraisals. Each student can then reflect on his or her own learning, time management, level of commitment etc., rather than simply placing the onus for learning on the shoulders of those who teach. There is a clear distinction between the processes of teaching and learning. You cannot evaluate the efforts made by one party without looking at the efforts made by the other party in my opinion because to do so results in an asymmetrical, unbalanced view of what's really going on. I teach an online paper and it records automatically the amount of time that students spend on the tasks they are required to undertake. I, therefore, get some sense of the amount of time and effort they put into their learning. I doubt many other courses offer faculty this insight. (Q21, sub-theme 2e, WU, junior teaching position, 11–15 years' tertiary teaching experience, continuing position, Commerce.)

I believe that students are over-evaluated, and sometimes evaluations are done when students have completed courses and many courses are evaluated at one time, which

can mean students become tired of filling out forms, get confused about courses and lecturers and so offer little information of value. I would like to see a systematic approach where students are informed of when evaluations will take place, what is being evaluated and why at this time. (Q21, sub-theme 2e, OP, lecturer teaching position, tertiary teaching experience not indicated, permanent position, Health Sciences.)

Under Theme 3 (Administrative changes), the issue of timing (Q21, sub-theme 3c, n=16, one per cent of respondents) is addressed, by the suggestion that *evaluations/appraisals should be carried out earlier so changes can be made and to avoid assessment activities which could influence data*. Some examples of the comments fitting under this sub-theme include:

[Evaluation] needs to be more qualitative, needs to be done BEFORE the course starts and at end to compare what has happened in between, needs to be followed up on the evaluations i.e. someone from HEDC go over the evaluations and discuss with staff. (Q21, sub-theme 3c, OU, lecturer teaching position, 21+ years' tertiary teaching experience, permanent position, Humanities.)

We should do appraisals the year after the paper is taken so that the question of how suited the paper was to future study in the same subject, or another subject which had the paper as a prerequisite, can be answered. (Q21, sub-theme 3c, WU, position not specified, 21+ years' tertiary teaching experience, continuing position, Humanities.)

Each subject area within a course should be evaluated at least once throughout the course. It is important to get the timing right, as students will not provide meaningful feedback if they are required to undertake too many evaluations at any time. In particular – timing to be outside that of First Impressions and Course completion – Polytechnic feedback. (Q21, sub-theme 3c, OP, senior teaching position, 11–15 years' tertiary teaching experience, permanent position, Commerce.)

In response to Q22 (General comments), a theme was identified around negative aspects/improvements needed (Q22, Theme 2, n=275, 51 per cent of respondents). Some comments under this theme identified the need for *better guidance on interpretation of results and how to follow up with students* (Q22, sub-theme 2d, n =24, four per cent of respondents), for example:

I think too much emphasis is given to the results. I have seen the results being used as evidence about how good or bad one's teaching is. And I am convinced that the evaluation results do not reveal that that clearly. To start with, the results are compared across many different papers that may be so different from each other in many ways. There is no recognition that some papers are difficult to teach. There is no recognition that the students may be rating a lecturer just based on how much they like the subject, rather than how good the lecturer is at teaching it. (Q22, sub-theme 2d, OU, lecturer teaching position, 16–20 years' tertiary teaching experience, permanent position, Commerce.)

I think school/faculty leaders should be more upfront as to how critical the role of the appraisals is in the promotions round. There should be some kind of benchmark perhaps (hard to do I know). But I am disappointed when I receive a 1.4 from an appraisal for teaching yet am told that others around the university are happy with a 2.7. Do we need a moderation process – that sounds rather scary to me but people need to know what's okay...I think it needs to be clarified – is this process summative or formative? This survey suggests it is both but that creates a tension. (Q22, sub-theme 2d, WU, lecturer teaching position, 11–15 years' tertiary teaching experience, continuing position, Humanities.)

Weighting needs to be attached. Who knows best what is in the students' best interest? A combination of them and the teacher, surely. (Q22, sub-theme 2d, OP, lecturer teaching position, 11–15 years' tertiary teaching experience, permanent position, no division indicated.)

Other comments related to timing (Q22, sub-theme 2i, n=11, two per cent of respondents) suggested *guidance on the use of evaluation other than at the end of semester*, as in:

I do not give students feedback on the outcome of evaluations as the formal evaluation process only occurs at the end of the paper at which point I have no further contact with students. (Q22, sub-theme 2i, OU, teaching position not specified, 0–5 years' tertiary teaching experience, fixed term position, Health Sciences.)

They should be given an appraisal at the beginning of the course listing their expectations etc. then perhaps in the middle of the course, and the final appraisal

should be a reflection of the whole semester. (Q22, sub-theme 2i, WU, junior teaching position, 11–15 years' tertiary teaching experience, position not specified, Commerce.)

The timing of evaluations are also critical. Teaching Unit standards means there are many times the evaluations are completed after an assessment. If a student is to show competency they may have 're-sits' to complete and this is not a good time for an evaluation but I am not sure when is the best time! (Q22, sub-theme 2i, OP, lecturer teaching position, 21+ years' tertiary teaching, permanent position, Commerce.)

d) How do lecturers use their evaluations/appraisals for professional development?

Guidance around the use of evaluation/appraisal data needs to be addressed through a range of professional development activities that will allow staff to actively engage with the process (Smith, 2008). The data gathered in the current study were examined to look for indications of whether teachers use evaluations/appraisals to determine professional development requirements, if at all. Question 2 (*Please identify why you use student evaluations/appraisals*) addressed this issue directly (see Table 4:22). In answer to Q2b (*For my professional development*) 87 per cent (n=793) of all respondents across all institutions identified this as a reason for using evaluations/appraisals, with OU ranking this as the third most important use at 88 per cent (n=499). OP ranked this as fourth out of the eight uses but was still high with 88 per cent (n=115) of staff using evaluations for professional development. WU rated this reason at a lower level (fifth, n=179, 83 per cent). What is very revealing is that the use for professional goal setting/salary review/confirmation requirements is ranked even lower at sixth (n= 544, 66 per cent) for all the institutions in the study. At WU the professional goal setting meeting is the official university mechanism for identifying professional development needs around all aspects of an academic's role (see section 2.3.1.2.4). At OU, the confirmation process is for staff new to the university and so it is unsurprising that this ranked lower as it would not apply to all respondents, for example, those in 'junior teaching positions', non-permanent staff and those who commenced their employment before the confirmation process was introduced. Information about this process is in 2.3.1.1.4.

As already discussed, Table 4:23 and Table 4:24 showed that slightly less than half of teachers actively discussed evaluation/appraisal results with colleagues or their teaching team (Q4e, 47 per cent). Thematic analysis of the comments associated with this question (Q5 *If there are other things you do with your student evaluation/appraisal results, please outline them here*) identified very few occurrences of interventions for professional development, and those that did, mentioned discussions with the teaching team in particular. The focus of most of the comments was on paper/course refinements.

Many of the responses to the comments questions in Section B of the questionnaire also mention professional development. For example, Q9 asked, *to what extent do your reasons for using student evaluations/appraisals influence your teaching decisions?* and comments to Q10 *please explain your answer to Q9*, which described reasons for influencing teaching decisions, Q10 sub-theme 1b highlighted that evaluations/appraisals enabled the development of the teacher and paper/course (n=301, 36 per cent). The details of the analysis of Q10 comments appear in Appendix 6. Comments in this sub-theme (Q10, sub-theme 1b) also mentioned identifying *teacher's strengths and weaknesses, uses for professional goal setting, course design, planning and refinement*. For example:

As the coordinator of two papers, I find the course evaluations to be especially helpful for planning the courses including methods of teaching (e.g. having a lot of class discussion) and assessment (refining the types of assignments). (Q10, sub-theme 1b, OU, senior teaching position, 11–15 years' tertiary teaching experience, permanent position, Health Sciences.)

I predominantly try to understand what things seemed to go well and what didn't. If I have made changes in my teaching or assessment I try to understand if these changes have had the desired impact. I particularly take note of the open question responses in this regard, because the big issues usually come through here. (Q10, sub-theme 1b, WU, lecturer teaching position, 16–20 years' tertiary teaching experience, continuing position, Sciences.)

If it is possible to do what students suggest re how the material is presented I usually do it. I have learned over the years though that each class is different and what is suggested by one class may not work with the next class, but over time good

suggestions become incorporated into my teaching. (Q10, sub-theme 1b, OP, lecturer teaching position, 11–15 years' tertiary teaching experience, position not indicated, Health Sciences.)

Similarly, responses to Q12 identified a number of themes around professional development (see Appendix 7). In answer to the question, Q12 *to what extent does your institution's use of student evaluations/appraisal data influence your teaching decisions?*, under Theme 2 *Developmental Purposes* (n=156, 21 per cent), the comments were clustered predominantly around one specific area, personal and paper/course development (Q12, sub-theme 2a, n=148, 20 per cent) with references being made to *personal/professional and course development and improving learning outcomes for students* (see Appendix 7). Thirty one per cent of responses were from OP staff, with 22 per cent and 16 per cent of respondents from OU and the WU respectively. A second, much smaller cluster of comments about professional development appears within sub-theme 2b, *alignment with broad teaching practices* (Q12, sub-theme 2b, n=8, one per cent) and covers comments about *enabling alignment with broader teaching practices, for example, colleagues, matching with graduate profiles and aiding development at departmental level*. These comments were from OU and WU only. Two examples are:

It is important for me to follow the strategies of my institution and my Department, so I take into account this information when I prepare my papers. (Q12, sub-theme 2b, OU, lecturer teaching position, 6–10 years' tertiary teaching experience, permanent position, Humanities.)

I need my teaching to be in line with the philosophies and practices of my teaching colleagues. (Q12, sub-theme 2b, WU, junior teaching position, 6–10 years' tertiary teaching experience, continuing position, Humanities.)

Another theme identified from the analysis of Q12 comments concerns promotion/quality control purposes (Q12, Theme 1, n=188, n=26 per cent). Here professional development is mentioned in a number of different comments, for example, under sub-theme 1b, *tension with institution use* (Q12, sub-theme 1b, n=73, 10 per cent of respondents). These comments indicate respondents' sense that institutional use of student evaluation/appraisal data *conflicts with formative use for own development or improving student learning*. What is interesting to

note is that these responses come from the universities only, both having 11 per cent of respondents (OU, n=51 and WU, n=22) making these comments. There were no comments in this category from OP. Examples include:

With the great deal of emphasis put on student evaluations in our promotions and confirmation progression, I view it foolish to proceed in a manner that would not resonate well with the students thus leaving us open to receiving bad evaluations. That said; I still attempt to push on with 'unpopular' elements of my courses if I view it as integral to their academic development. (Q12, sub-theme 1b, OU, lecturer teaching position, 0–5 years' tertiary teaching experience, confirmation position, Humanities.)

The institution only uses the data for quantitative matters. More emphasis needs to come on how individuals can use the appraisal system to improve and develop their teaching. Currently, appraisals are compulsory, but interpreting them is not compulsory so teachers tend to justify particular appraisals rather than trying to interpret the results. (Q12, sub-theme 1b, WU, lecturer teaching position, 6–10 years' tertiary teaching experience, continuing position, Commerce.)

Similarly, this pattern is repeated for sub-theme 1c, *institution does not value or use the information* (Q12, sub-theme 1c, n=23, three per cent of respondents). Comments here show *little interaction or follow-up when needed* identifying the perceived lack of follow-up processes around poor teaching as identified by evaluations/appraisals. Again, these comments were made solely by teachers from the universities. For example:

The University uses student evaluations for staff performance reviews, but I am not aware that evaluations are followed up if problems are identified. (Q12, sub-theme 1c, OU, lecturer teaching position, 21+ years' tertiary teaching experience, permanent position, Health Sciences.)

I have found over time that there does not seem to be any real institutional interest in the results. It appears to be more a box ticking exercise. (Q12, sub-theme 1c, WU, lecturer teaching position, 21+ years' tertiary teaching experience, continuing position, Sciences.)

In terms of professional development, a number of improvements were suggested, in responses to Q21 (see Appendix 12 and Appendix 13). As already stated, Q21 asked staff about their views on the future of evaluation/appraisal at their institution. Just over a third of respondents' comments for this question clustered under Theme 2 and had a *Developmental Focus* (Q21, Theme 2, n=363, 32 per cent). Under this theme are a number of sub-themes that mention development. The predominant sub-themes of Theme 2 are 2a concerning the need for *Flexibility* (Q21, sub-theme 2a, n=99, nine per cent) and 2d concerning the *use of other evaluation methods* (Q21, sub-theme 2d, n=104, nine per cent). Sub-theme 2a addresses the need for flexibility with evaluation instruments, both in format and questions and to allow teachers to design their own evaluations and ask specific/appropriate questions. The three institutions were fairly evenly matched with eight per cent for OU (n=57), eight per cent for WU (n=24) and 11 per cent for OP (n=18).

I prefer a more local format; targeted specifically to what I really need/want to know about my teaching or the local course. I would prefer not to have the compulsory questions as they're more about how much the student likes you than how effective you are. (Q21, sub-theme 2a, OU, junior teaching position, 6–10 years' tertiary teaching experience, permanent position, Health Sciences.)

Keep the centralised system for processing surveys – it is very efficient and done in a very helpful way. But allow all of the questions in questionnaires to be generated by the teachers of each course so that they are appropriate for the course and can seek feedback where it is particularly needed. (Q21, sub-theme 2a, WU, senior teaching position, 21+ years' tertiary teaching experience, continuing position, Sciences.)

Focus on learning outcome results. (Q21, sub-theme 2a, OP, lecturer teaching position, 11–15 years' tertiary teaching experience, permanent position, Division not indicated.)

Comments collected under sub-theme 2d mention that *there should be more integration with other evaluation methods, that other methods are better for development and to assess quality, for example, peer and expert review, class reps, informal dialogue, informal surveys, quizzes, focus groups and so on*. These comments came mostly from respondents at the

universities, with 11 per cent (n=32) of the responses from WU, nine per cent (n=62) from OU and six per cent (n=10) from OP respondents.

I would look for more flexibility in the format and the questions to suit different teaching situations. I would encourage other forms of evaluation (peer etc.) to be rated as equally if not more important than instant customer feedback. (Q21, sub-theme 2d, OU, senior teaching position, 21+ years' tertiary teaching experience, permanent position, Humanities.)

If it were the only tool used to evaluate a teacher's ability then I would definitely not be in agreement with it. Used along with other forms of appraisal, it is a very valuable tool. (Q21, sub-theme 2d, WU, teaching position not indicated, 0–5 years' tertiary teaching experience, fixed term position, Humanities.)

I would abandon very formal evaluation. For refining and developing my curricula I prefer an informal and dynamic evaluation process. (Q21, sub-theme 2d, OP, lecturer teaching position, 21+ years' tertiary teaching experience, permanent position, Humanities.)

Comments under sub-theme 2b, *Developmental use* (Q21, sub-theme 2b, n=38, three per cent) also say that there should be *more emphasis on developmental use and course appraisal/evaluation for formative purposes*. For example:

More emphasis on the evaluation of the design, delivery, and effectiveness of whole papers, rather than just the teaching of individual lecturers. (Q21, sub-theme 2b, OU, senior teaching position, 21+ years' tertiary teaching experience, permanent position, Humanities.)

I think the appraisal process should continue but its primary aim should be for staff professional development to improve learning outcomes for students. I would like it to be incorporated further into teaching development and valued more highly in terms of promotion, particularly for teaching only staff. (Q21, sub-theme 2b, WU, junior teaching staff, 16–20 years' tertiary teaching experience, continuing position, Humanities.)

Combine the process to include both course and teacher evaluation, to simplify it for staff and students. (Q21, sub-theme 2b, OP, lecturer teaching position, 6–10 years' tertiary teaching experience, permanent position, Humanities.)

Guidance/support of teachers is also a sub-theme that was identified (Q21, sub-theme 2c, n=24, two per cent), with suggestions including *how to use as a tool for improvement (i.e., educate the staff in formative uses), how to interpret and act on the results*. Examples of comments categorised within this sub-theme include:

I think a more descriptive analysis of the data would be helpful, from an external reviewer with an educational background, who might offer suggestions on how to address the evaluation and share ideas about making useful changes, using educational principles. (Q21, sub-theme 2c, OU, lecturer teaching position, 11–15 years' tertiary teaching experience, confirmation position, Health Sciences.)

A more collegial/academic mentor and 'professional development' model needs devising, in which student feedback is constructively sought from which teachers would be taught about interpreting the data and what modifications to make to their courses in response. (Q21, sub-theme 2c, WU, junior teaching position, 6–10 years' tertiary teaching experience, continuing position, Humanities.)

The one thing I would do is include evaluation education sessions in our whole staff training days. I don't think all staff know how to access the evaluations or that they can modify them to evaluate specific things. I think it would also be good to educate staff about the professional development aspects of evaluations as I think lots of people perceive them as punitive things. I think that any punitive aspects should be played down – not helpful for anyone! (Q21, sub-theme 2c, OP, senior teaching position, 0–5 years' tertiary teaching experience, permanent position, Humanities.)

4.3.5.2 Teacher engagement in evaluation – evidence from the interviews

4.3.5.2.1 University of Otago

There is a considerable range in the ways in which OU interviewees spoke about how they engage with evaluation results. A significant number of respondents indicated that they use results to inform their teaching, although this process varies from deliberate and systematic

usage to a more cursory 'nod' to the student feedback. Five interviewees spoke about deliberate and relatively systematic responses they make to the student evaluation feedback they gathered. Examples of comments include:

Uses them to inform practice. Systematically collects many sources of information to review. Uses them in planning.

Undertakes to read them all and make changes if necessary (but uses multiple sources.

Has a system of identifying themes from evaluations. Doesn't find ratings useful on their own.

Looks at comments to improve teaching.

Has a systematic process for going over the evaluations and will attend to new ideas that a number raise.

Six respondents indicated that they consider student feedback on their teaching to some extent and for varied purposes. Comments include:

Makes use of comments to see if new approaches are working. Thinks about them. No actions taken.

Main aim is to learn from evaluation so as to get a good score next time.

Looks at them reflectively and in comparison with past results...it is just one little thing in amongst the huge amounts of teaching and administration and everything else I'm trying to do.

Wants to achieve good appraisals, but not to the extent of buttering them up.

The rest of the respondents indicated a lack of engagement with evaluations/appraisals feedback, which ranged from apathy to negativity. Comments include:

Needs encouragement to go over the results.

I've long gotten over being surprised by mainly this point about people liking and loathing the same thing. So now actually, I don't attend to them.

Doesn't use them to improve. Nothing very valuable in them. Vacuous. Written in ten minutes what you have thought about for years.

Definitely feels constrained from trying new approaches.

There was no strong evidence of a pattern of closing the loop around evaluations with very patchy reference to discussion with colleagues and feedback to students. Seven interviewees reported some degree of discussion with colleagues, but that this is not always systematic.

Comments include:

Informal meeting/discussion with colleagues when looking at the programme as a whole.

Sometimes shares information with other teachers

– only informally.

Has conversations with others.

Six interviewees said they discussed evaluations with their students, ranging from deliberate purposeful communication to informal chats. Comments include:

Reports back to students on any changes made in response to previous evaluations.

Important that students believe the process is worthwhile. And that involves them knowing why they are doing it, what might happen as a result of it, and what that might mean to them.

Talks to students about the uses of evaluations.

Feedback to students both directly and indirectly.

Informal conversations with students.

Does let students know about changes in response to evaluations but not systematically or regularly.

4.3.5.2.2 The University of Waikato

WU respondents varied considerably in the degree to which they claimed they engaged with student feedback from student appraisals. As at OU, there is a small group who said they use

the student evaluations feedback in a deliberate and systematic way to inform their teaching (four). Comments from educators in this group include:

I go back to my reflective journal and I have a look at my thinking about their learning. Sometimes, you'll have a group of people who make cognitive comments, so I take that very seriously...I apply them to the planning of my paper for the following year.

What I do is look at the comments, and try to adjust in response to the comments. The grading system is really useful because they tell me things are on track, whereas the comments give me direction about what could be improved.

I type up a list of the general comments that students make and act on those if I can... I'll often ask them to comment on things I'm trying or testing out, and I want to know if they're working.

I take their comments and use them as objectives for myself of things that I need to change and adjust...I ask myself, how can I transform this into a teaching objective?

Ten WU respondents indicated that they engaged to some extent with student feedback. Their comments include:

If several students come up with the same idea, then I try to incorporate it or if one student comes up with a brilliant idea, I try to incorporate it.

I'm more concerned with their written responses. When that comes through I like to see if there's anything I can do differently, what they liked, disliked.

If it's negative then I change my course...If there's a hint of a student in trouble. Then I'll follow it up in various ways.

I primarily use them as part of the review process.

Three respondents at WU said that they do not engage with evaluation feedback at all, while the remaining three WU interviewees did not discredit the notion of using the feedback, but offered no elaboration as to what they did or might do with this information.

The interviews at WU did not reveal a significant culture around discussion with colleagues about student evaluation feedback and ways of responding to or utilising it. It would appear

that peer conversation around student evaluations is not a normative and expected part of the culture. While ten interviewees indicated that they do talk to colleagues about student evaluations/appraisals, this is not usually reported as being a deliberate and systematic process. Two of the ten said they do look regularly at student evaluations/appraisals with colleagues as part of the paper review process. The others who said 'yes' to this question used language that suggests that the process is informal, non-specific, unstructured and, in one instance, almost 'flippant'. Comments include:

But I tend to say the feedback was better or worse than last year, don't go into specifics that much.

Sometimes/informal/in the staffroom but not in meetings.

Yes just what strategies they're taking.

Yes, we do discuss but tend to be a bit flippant about it...It's a throw your hands up in the air sort of topic.

Although seven of the interviewees gave an unqualified 'no' to the idea of discussing evaluation feedback with students, there is some indication that WU interviewees recognise the value of discussing feedback with students. Four interviewees said that while they do not discuss end of semester feedback with students, they obtain feedback from students during the semester, and they indicated that they discuss these in-semester findings with students, making modifications in response to the feedback, where appropriate. Comments include:

I do discuss mid-semester informal appraisals. I present it to them graphed and categorised and say what we will do about it. I think that's most valuable because of the time it comes.

I get ongoing feedback just in the normal course of a lecture... I make it clear to them that we do value feedback, that we're looking at ways to evaluate and improve their learning.

Seven WU interview participants said that they do not discuss the feedback and their responses with the current group of students, but the majority of this group (six) noted that the discussion happens with the next cohort of students. For example,

If there's stuff that's interesting and relevant, I usually discuss with incoming classes what I've learned from previous classes.

The remaining two WU participants indicated that they only discussed feedback with a particular student under specific circumstances, such as when a student has raised an issue with them directly.

4.3.5.2.3 Otago Polytechnic

There is a clear difference in the extent to which OP interviewees say they engage with student feedback as compared with respondents from the two universities. Seventeen reported that they took account of the evaluation feedback provided by students and used it to modify their teaching. While the degree of engagement with student feedback and its incorporation into changes in course design and delivery varied considerably across the OP interviewees, there was a general recognition of this process as a necessary and constructive routine.

Examples of comments include:

I use feedback to look for themes about learning styles and teaching methods. I use feedback to think about and adjust teaching to engage students in learning.

I read the scales and then the comments. I do respond to useful suggestions.

They are really helpful...useful...I actually will adapt my classes early on to meet as many of these styles that I can.

I see it as quite a continual process...if they are done earlier then there is better feedback and understanding between the students and the teacher.

In spite of the high number of interviewees who reported that they found student feedback useful for their teaching (17), only five of the OP interviewees reported discussing their response to student feedback with the students. This relatively small number of academics who saw it as important to complete the feedback loop in this way matched the small numbers also discussing feedback with students from the two universities. Furthermore, even those who did say they discuss their responses with students, did so primarily on the course evaluations, as opposed to the evaluation of teaching. Comments include:

I do occasionally tell them we are making these changes because of feedback.

I don't say what I've done with it to the students, just in a very general way.

If I get an evaluation that needs to be actioned, I will tell them this has occurred and this is what we are going to change.

On the whole I don't discuss them. The paper that I coordinate is in Semester 2. They have moved on to the next course.

As in the case of the universities, the timing of evaluations is cited as a significant reason for not communicating responses to feedback to students.

4.3.6 Conclusions

In conclusion, it appears that most teachers engage with the evaluations/appraisals process, but to varying degrees. Many believe that collecting evaluation/appraisal data is worthwhile, mainly for ongoing paper/course refinements, and to receive feedback on the students' learning experiences. However, there is a small number of staff who believe that the data that students provide can be biased, and based on poor judgement. There appears to be little feedback of evaluation information to students, mainly because of the timing of the evaluation/appraisal process, but many staff do indicate that they use other forms of evaluation throughout their teaching.

In terms of professional development, many teachers do not actively seek help with using evaluation/appraisal data, even though a small group of teachers have problems with interpretation of the data. From the qualitative questionnaire comments, issues were identified with follow-up professional development processes, educating and supporting students around the evaluations/appraisals process, and also staff engagement being reduced as a consequence of institutional restraints and requirements. It would appear that many teachers deal with their evaluation/appraisal results in a somewhat isolated way.

What is quite clear is the difference in the way that teachers from OP and the universities engage with the evaluations/appraisals process. OP's use of evaluations/appraisals is part of a structured quality assurance process that focuses mainly on the provision of good teaching practice. The universities, meanwhile, have not had such well-defined structures or processes, because their focus is spread more widely across research, teaching and other such activities

that define academic activity (see Table 2:4). How this is likely to change in the future will be of interest for both types of institutions.

4.4 Summary of Chapter 4

Chapter 4 presented the findings of the study. These findings, around the three questions that arose from the research intentions, provide a basis from which a series of assertions about tertiary teachers' perceptions of student evaluations were developed. These assertions, which were devised as a way of expressing threads of meaning across the research questions, are discussed in Chapter 5.

Chapter 5 : Discussion

5.1 Introduction

In this chapter, in line with the interpretivist research approach that framed the project, the findings and outcomes of the study are now brought together. Because evidence is a means to be able to provide a reasonably confident set of claims about the phenomenon under investigation, no declarations are made about ‘what is’. Rather, the evidence is presented in a way to show what seems ‘highly likely to be the case’ (Miller & Fredericks, 2003).

At the end of Chapter 2, through a combination of recent research on evaluation and current practices made public through policies and processes of New Zealand universities and polytechnics, the breadth of the ‘territory’ under investigation was mapped out in a conceptual framework (see Figure 2:1). The overarching research question was: *How do the current evaluation processes and practices influence teachers’ thinking and behaviours in relation to student learning at all stages of the teaching and learning cycle?*

Contributing research questions were:

1. *What perceptions do tertiary teachers hold about student evaluations?*
2. *What factors (causes, influences) affect these views?*
3. *How do tertiary teachers engage with evaluation results and student feedback?*

Chapter 3 then described the research design and Chapter 4 presented the findings.

Based on those foundations, the discussion is now presented in the form of a series of assertions that are underpinned by a refined version of Figure 2:1, shown in Figure 5:1, below.

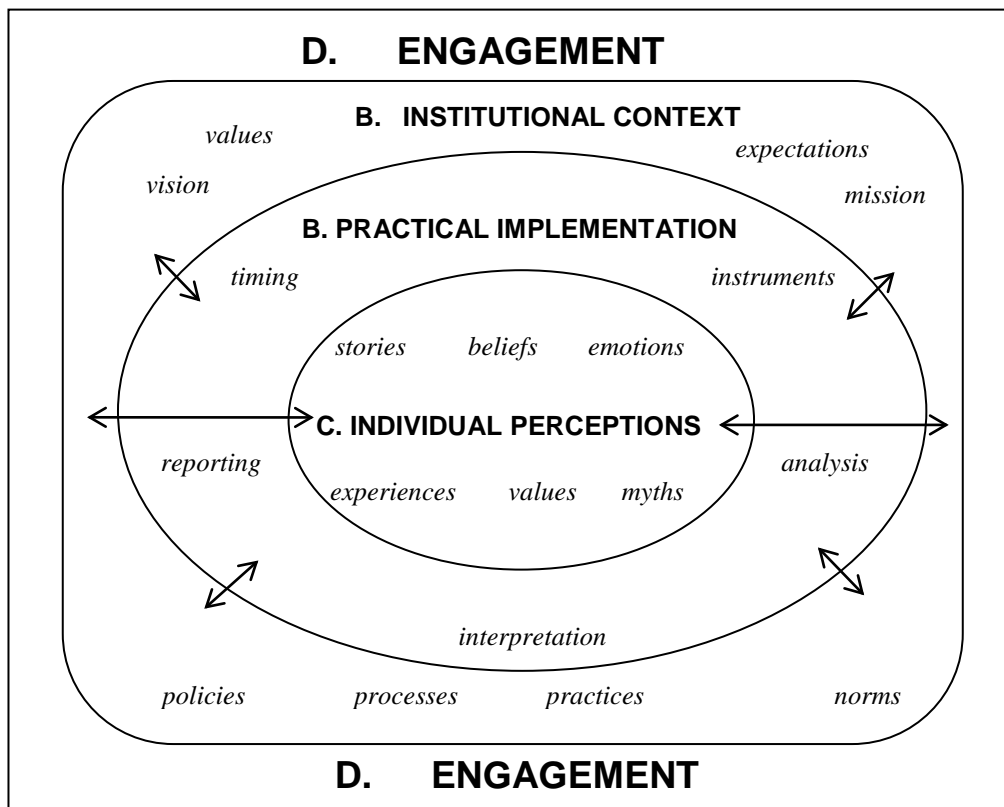


Figure 5:1 Factors influencing engagement with evaluation

Figure 5:1 shows refinements to the central section of Figure 2:1. It brings into focus the many factors influencing engagement in student evaluation/appraisal of teaching and courses that emerged in the light of the data analysis and findings from this study.

It was shown in Figure 2:1 that perceptions are influenced by attitudes and beliefs as well as by contextual factors. An individual teacher will have beliefs, understandings and perceptions about learning and teaching, about students, and about the role and place of evaluation, and that individual will live and work among others who also have beliefs, understandings and perceptions about teaching, learning and evaluation. As a member of various communities – institutional, discipline, departmental/school – that individual will be subject to community demands and expectations. The formal policy, or ‘official word’, exists alongside the less formal ‘translation’ or enactment of the official word, operationalised then experienced through various practices, systems, processes, procedures and norms.

An individual’s perceptions of student evaluations are determined not only by already held belief and attitudes, but also shaped and influenced by those of the community members with

whom the individual works and interacts, and the contexts in which the community operates, including the various practices, processes and procedures that must be engaged with in order to contribute and participate in that community. The individual is therefore subject to:

- their own and others' stories, myths, emotions and experiences related to evaluation (the *individual perceptions* section of Figure 5:1)
- the necessity to utilise, and make sense of, the ways their institution operationalises its policies around evaluation (the *practical implementation* section of Figure 5:1)
- the wider realm of institutional expectations (the *institutional context* section of Figure 5:1).

The following composite case studies with accompanying implications for future practice and research were assembled using the data from the interviews as a way of capturing the 'typical' sets of views expressed across the interview groups from each of the institutions. The case studies serve to highlight the range of views and how those views are revealed in the complex practices and behaviours of the teacher. The case studies also highlight that contextual factors play a part in shaping and informing views about teaching, learning and evaluation. They thus serve to provide an illustration of the individual dealing with practical implementation within the wider institutional context (Figure 5:1).

OU Case Study 1 – Miriam

Miriam is a health professional who started teaching at university in her forties. She has no background in teaching and still conceptualises herself primarily as a health professional. She is keen to give the students the benefit of her clinical experience but feels that students do not really appreciate the requirements of the profession. For this reason, she believes that the students are fairly poor evaluators of teaching and learning, and she cannot help feeling personally hurt by some of their comments. However, Miriam is apprehensive about student evaluation results as she knows that they may influence her promotion chances. Lately, she has been selecting questions that seem safe and on which she usually achieves a solid score. She does not discuss her evaluations with colleagues or students because she believes that they are private. In any case, she does not believe that she has time to spend on evaluation 'post-mortems' as she is under pressure to build her research profile.

OU Case Study 1 – Miriam – Implications for future practice and research

Miriam's case highlights the needs and challenges of a particular cohort of higher education teachers, namely, those who come into tertiary teaching on the basis of their expertise in a profession or trade. This situation was particularly noticeable from the interviews at OU, where there is a strong emphasis on professions such as medicine and dentistry, and at OP, an institution which is highly focussed on preparing students for employment in a wide range of fields. The needs of this group suggest that institutional requirements of teacher and professional development should specifically help staff to translate their industry/professional knowledge and practice into forms and approaches that are suitable for formal teaching, learning and evaluation contexts.

While Miriam's case highlights some of the challenges of moving from the workplace into teaching, her comments also pinpoint another barrier to the proactive use of evaluations/appraisals to enhance teaching in the university sector, that is, the imbalance between the regard given to research and that given to teaching. For some teachers who participated in this study, a preoccupation with research was evident in their expressions of their teaching philosophies. This preoccupation was also evident in data concerning activities teachers engaged in when they received student evaluation/appraisal results, implying that it is not worth engaging too energetically with student feedback on teaching. Miriam's case also highlights more general themes that emerged in the study, such as the mistrust of students' capacity to judge teaching and the fear of institutional reprisals if appraisals are perceived to be poor. Her emotional concerns highlight the need for further research on the affective dimensions of student evaluations/appraisals.

Universities, in particular, need to demonstrate in practice, that the institution gives identical weighting to teaching and research in promotion decisions, and be transparent about how promotion decisions are made. This may then help to counter the supremacy that research has been seen historically to have over teaching in universities. Additionally, it needs to be a requirement that the majority of academics who enter the university on the basis of their discipline research expertise, engage in professional development around teaching and learning theory and practices.

OU Case Study 2 – Joshua

Joshua is a Humanities teacher who wants to develop students' ability to think for themselves and to help them to develop habits of effective and purposeful questioning. He is always interested in the student evaluation feedback, although does sometimes have the sneaking suspicion that lecturer popularity is more important in student evaluations than it should be. At the beginning of the semester he talks to the new cohort of students about how he responded to the feedback from the students in the previous year and explains the changes that he has made. He also explains the importance of their feedback for his teaching. During the semester, Joshua conducts his own informal student evaluation exercises every three weeks and then discusses his response with his students. Sometimes, he revisits a topic or point if the students indicate they do not understand it. Joshua is always attentive to the feedback in the formal student evaluations, but generally finds that there are few surprises for him because he elicits regular feedback. He is not worried about the university's use of student evaluations for promotion, as he sees them as only one source of information about his teaching.

OU Case Study 2 – Joshua – Implications for future practice and research

Joshua's case exemplifies a particular perception of the role of tertiary education that was shared by some of the interviewees. Respondents like Joshua emphasised the formative potential of higher education and focussed strongly on ways of developing students' capacity to inquire in their discipline. In the study, there was a noticeable correlation between these kinds of perceptions and teachers' interest in student feedback, as well as an interest in engaging in ongoing dialogue with their students. The case of Joshua, and others like him, indicates that how to engage meaningfully with student evaluation/appraisal and their integration into the teaching and learning cycle is part of a more general need for professional development in teaching and learning.

Joshua's case also demonstrates how the use of multiple forms of ongoing evaluation can promote dialogue with students and a raised awareness of their needs. His case suggests that institutional endorsement and encouragement of a range of student feedback strategies can help to make evaluation activity a central, integral and dynamic part of the teaching and learning process.

WU Case Study 1 – Jenny

Jenny is an experienced academic in the Social Sciences. She believes that her primary role as an educator is to enable the students to acquire the capacity to challenge and question the communities in which they will participate at different stages of their lives. She wants her students to become critical professionals. She values students' feedback but feels that the university does not effectively encourage staff to use appraisals for teaching and professional development. Consequently, she believes that staff members have become cynical about the appraisals process. She is annoyed by the timing of the student appraisals, because it means that she cannot make improvements for the current cohort of students. Jenny collects regular informal feedback from her students and tends to use the university system just to satisfy institutional requirements or for promotion. She talks to students about her response to their informal feedback, but does not talk with her colleagues as they already have the opinion that she spends too much time and energy on teaching as opposed to research.

WU Case Study 1 – Jenny – Implications for future practice and research

Jenny's example, like Joshua's, illustrates an academic with a strong focus on nurturing thinking and inquiry in students. In Jenny's case, this is associated with a commitment to developing students' capacity to critique established norms, practices and beliefs. Jenny's focus on critical inquiry and challenge overrides transmission of content as a teaching and learning priority. Jenny's sustained interest in students' thinking means that she gathers feedback in a variety of informal ways across the semester, seeing the formal appraisal as simply an institutional requirement. Furthermore, the potential developmental benefits that the formal system might be able to facilitate are undermined – in Jenny's view, and that of many academics in this study – by the timing of formal appraisals. The inappropriate timing of student evaluation/appraisal is clearly an issue that institutions need to address. Jenny's preference for ongoing student feedback suggests that the institution needs to recognise and acknowledge the place and value of multiple forms of evaluation feedback in its promotion processes.

Jenny's reference to not talking to her colleagues represents a more widely held assumption about privacy around evaluation/appraisal feedback, as well as a common perception that research is more valued than teaching at universities. Views of this nature indicate that, to ensure there is a match between claims made about teaching in vision statements and policy documents, universities need to demonstrate that through actual practices, all matters around

teaching and teaching performance are applied with the same degree of seriousness that is attached to research.

WU Case Study 2 – Paul

Paul has been teaching in the Sciences for 30 years. He thinks teaching is important and aims to explain core concepts to his students in a clear and accessible manner. He always tries to find fresh ways of making the material engaging for his students. He is interested in student feedback gathered through formal evaluation, but feels that often, students cannot make objective judgements because their understanding of the subject is inevitably partial. He does not refer to student appraisals to inform his teaching unless he gets an unusually low score. In that instance, he may go back to the comments to try to find out what has happened. At the same time, he is not keen on ‘kneejerk reactions’ in response to student appraisals as there is a curriculum that must be covered. Paul does not talk with colleagues or students about his evaluations.

WU Case Study 2 – Paul – Implications for future practice and research

Paul’s teaching beliefs exemplify a way of thinking about teaching that this study confirmed is still highly visible in the universities: that teaching, essentially, involves the transmission of a body of knowledge, as clearly and efficiently as possible. Generally, teachers who expressed this view of teaching in this study were less responsive to student feedback gathered through evaluations/appraisals than their counterparts who held more student-focussed views of teaching, as their commitment, they claimed, is first and foremost to the delivery of the fixed body of knowledge. Paul’s views indicate that attitudes to student evaluation/appraisal need to be seen within the context of a more general study of teaching and learning. Paul’s case also illustrates the need for institutions to address the professional development issues of longstanding academics as well as new ones.

OP Case Study 1 – Jim

Jim is an electrician by trade and has been teaching at the OP for the last 10 years. He teaches at all levels of the Certificate and works side by side with his students on a regular basis. He sees his educational role as a trainer who is sharing from his own knowledge and experience

and preparing students for their vocation. He believes that students can recognise a good learning experience, but sometimes are swayed simply by whether they like a teacher or not. He also thinks that some students may take it out on the teacher if they do not pass the assessments. He reluctantly admits that he occasionally takes negative feedback a bit personally. He also believes that some students are too immature to give feedback, especially if they are just out of school and lack the vocabulary and capacity to articulate their views. Jim is not overly focussed on the formal student evaluation feedback because he works alongside students all the time, so regularly finds out about their understanding and thinking. He is not keen on discussing students' comments directly with them, although sometimes he will tell them if he has made changes in the light of their feedback.

OP Case Study 1 – Jim – Implications for future practice and research

Jim's case illustrates some of the contextual factors that are pertinent to the way teachers regard feedback from formal student evaluations at OP. In particular, his comments include reference to student groups at Certificate level, who, he believes, lack the maturity to make dispassionate judgements about teaching. This view of his students is accompanied by a defensive and strongly emotional perspective on student evaluation/appraisal in general, alongside a belief that students tend to make judgements on the basis of factors such as teacher likeability.

While Jim's concerns about students' ability to formulate and articulate judgements are intensified by the fact that many of his students are young and have had limited previous educational opportunities, like many participants in the study, the mistrust of students' judgements was a pervasive theme across all three institutions. This finding suggests that institutions, academics and those charged with academic staff development and student evaluation/appraisal, need to improve education around the rationale for, and benefits of, student evaluation/appraisal. Correspondingly, institutions need to set up policies and systems to ensure that these benefits are made transparent to students.

OP Case Study 2 – Mere

Mere is an educator on a degree programme at the OP. She sees her role as prompting students to think about, and engage with, social justice issues. She talks about transformative learning and her hope is that the learning experiences she provides will be transformative for

her students. She is an avid collector of student feedback and is committed to closing the feedback loop. She believes that students need to be listened to and shown that their views matter. She does have some concerns about the quality and usefulness of the questions on the standard formal evaluation questionnaires and tries to collect feedback throughout the course and discuss it with students. She feels that the institution is too focussed on the quality dimensions of the evaluation and that it does not promote and support the professional development benefits of the instrument strongly enough.

OP Case Study 2 – Mere – Implications for future practice and research

Mere's case illustrates a connection between a transformative vision of teaching and learning and a sustained interest in student reports about their learning progress and experiences. Like a number of her counterparts, Mere uses ongoing formative evaluation to gauge her students' learning and believes that the institution has a role in highlighting and proselytising the educational development benefits of student evaluations. It would be important for the institution to provide flexible instruments so that academics such as Mere could use them as practical tools that recognise the importance that evaluation has on learning. Similarly, the inclusion of a core, standard set of meaningful questions within a flexible instrument would support individual teacher and institution needs to monitor quality over time.

In the current, high accountability agenda context, institutions are under pressure to provide evidence of quality monitoring. However, unless monitoring and auditing effectiveness and performance are accompanied by a visible emphasis on evaluation/appraisal for development, institutions run the risk of fostering cynicism and disengagement by academics. Too unbalanced an emphasis on quality at the expense of development will not improve engagement with student evaluation/appraisal to enhance student learning, but will increase academics' tendency to see and use them as an isolated add-on to their work and nothing more than a required ritual.

5.2 The Assertions

Assertions, listed in Table 5:1 below, were developed about the evidence that was gathered through this study to highlight the outcomes through focussing on the interplay among the aspects shown in Figure 5:1, A (*individual perceptions*), B (*practical implementation*), and C

(*institutional context*) influencing D (*engagement*) in student evaluations, illustrated through the composite case studies in the previous section.

Table 5:1 Assertions

		Connections highlighted in Figure 5:1
Assertion 1	<i>Institutional policies, processes and procedures influence teachers' perceptions about student evaluation/appraisal and their associated behaviours</i>	C, A, D
Assertion 2	<i>Teachers' perceptions about the quality of student feedback including students' ability to make judgements about teaching and courses influence how teachers view student evaluations and engage with evaluation/appraisal data</i>	C, B
Assertion 3	<i>The use of the same instrument for professional and course development and for judging the quality and effectiveness can limit teachers' views of evaluation/appraisal and their evaluation-related behaviours</i>	B, D
Assertion 4	<i>Teachers' beliefs about teaching and learning and how they view themselves and their role in their institution will influence their perceptions of, and engagement with, student evaluation/appraisal</i>	C, A, D
Assertion 5	<i>Teachers tend to view student evaluation/appraisal as an isolated and individual activity that informs them about their courses and provides data to demonstrate their teaching effectiveness to their institutions</i>	C, D

The assertions are now discussed in the light of the literature. Explicit links between the assertions and the elements in Figure 5:1 are made.

5.2.1 Assertion 1

Assertion 1: *Institutional policies, processes and procedures influence teachers' perceptions about student evaluation/appraisal and their associated behaviours.*

This assertion concerns how institutional factors influence teachers' evaluation thinking and behaviours. The assertion focuses principally on the connections between section A in Figure

5:1 (*institutional contexts*) and section C (*individual perceptions*) and the impact of those connections on section D (*engagement*).

Institutional expectations versus individual perceptions

One of the first prompts that led to the instigation of this research project was the high degree of anxiety that academics often demonstrated when talking about student evaluations and especially the comment that the institutional use of student evaluations for quality control and promotion undermined its development possibilities. During the course of this investigation comments of this nature have been regularly expressed at Ako Aotearoa colloquia. It was the starting premise of the researchers that tension between the institutional use of evaluations for quality and promotion and its development role for academics was at the heart of the way academics perceived student evaluations. There was considerable evidence from the literature to strengthen this view.

The nature of teachers' engagement with the student evaluation system at their institution varies according to the purposes they believe the institution has for the system and how they react to the practical implementation aspects of their institution's evaluation/appraisal system. Mistrust or suspicion of institutional use of appraisals has been widely voiced in numerous studies over a long period of time (Arthur, 2009; Beran & Rokosh, 2009; Costin *et al.*, 1971; Edström, 2008; McKeachie, 1997; Moore & Kuol, 2005). The questionnaire findings in this research indicated that institutional use of evaluations/appraisals is a factor for some academics (see discussion in section 4.3.3.2, Table 4:16), but is not nearly as pervasive and dominant as was expected at the outset of the project. As shown in Table 4:12, 73 per cent of all participants said that they thought student evaluations were personally worthwhile. This result might mean that staff would be involved in student evaluations even without institutional directives.

Only two interviewees from OP explicitly raised the question of institutional uses of student evaluations. One of them saw it as a problem while the other expressed concern that the institution did not do enough to encourage teachers to use student evaluations for professional development. However, it should be noted that OP has two completely separate questionnaires, one for the course and one for the teacher, and that some respondents (five) indicated that they saw the course evaluations as more of a public document and the teacher

evaluations as a private document. The theme of suspicion of institutional use of student evaluations was much more evident in the interview findings from OU. Eight interviewees expressed in different ways the sentiment that the institution's use of student evaluation data limited its usefulness as a professional development tool. Frequently accompanying the perception that the institution's use of evaluations tainted their professional development value, was the idea that the use of evaluations for quality control and career advancement promoted manipulation of the questionnaires by academics. Eight interviewees voiced the opinion that the questionnaires were manipulated by teachers in order to engineer a better result. Like OP, OU has two different questionnaires, one for teaching and one that focuses on courses. However, both are possible inclusions in documentation required for confirmation and promotion purposes, although historically much more emphasis has been placed on the teaching evaluations. At WU, the interview findings did not reveal such strong signs of mistrust of institutional uses of student evaluations as at OU. Three WU interviewees voiced these concerns although three more argued that the university needed better systems in place to encourage and support the use of student evaluations as a developmental tool.

At OP and WU institutional expectations are made very clear through policy (see Table 2:5) with annual (OP) or biennial (WU) evaluations/appraisals being mandated. At OU, perhaps in line with its Value statement about 'academic freedom' (see Table 2:4), there are no mandates about when and how often evaluations are to be conducted, but to meet confirmation path and promotion requirements, teaching evaluation results are to be submitted. For OP participants, 97 per cent said that they use student evaluations because is a requirement of their institution (see Table 4:16). In addition, 53 per cent claimed that their teaching decisions were influenced by their institution's use of evaluations (*i.e.*, 80 per cent giving Q11 a rating of 1, 2, or 3 – see Table 4:17).

So, in essence, the evidence suggests that overall, a combination of factors causes teaching staff to run student evaluations/appraisals and to make use of the data. It appears to be a mix of seeing inherent value in being involved in evaluation (as a learning and reflective process – for example, see Table 4:13, Q2f, e and b) and the necessity to report on teaching and course effectiveness and quality to internal and external bodies and respond to the demands placed upon teachers by their institution (see Table 4:13 –for example, Q2a, c and h). Each of these factors is not mutually exclusive, of course.

According to studies cited in the literature review, feelings of suspicion and mistrust (for example, Beran & Rokosh, 2009; Burden, 2008; Moore & Kuol, 2005; Penny & Coe, 2004) can surround staff views of institutional use of student evaluation data. While there was a generally high positive view expressed about the worth of evaluations/appraisals in the current study, when other data were examined it was evident that those generally positive claims needed to be qualified.

The particular concern highlighted by Costin *et al.*,(1971) about doubts staff have regarding the way student evaluation/appraisal data are used by those who make decisions in institutions was also evident in data collected from the participants in the current study.

Doubts were expressed about the results of single survey instruments being used to make decisions about teachers' futures (see further discussion on this point in section 5.2.3). Such a focus indicates strongly that institutions assume student evaluation/appraisal results, on their own, can capture the complexities and nuances of teaching and the learning environment. Even in the case of OU where evaluation results are expected to be presented in a reflective statement, demonstrating teacher learning and development in the light of student evaluation results, (in the *Otago Teaching Profile* – see section 2.3.1.1.4 and Table 2:5) perceptions held by some staff who were interviewed were strongly negative about the institutional use of the data, as well as the potential for manipulation of the results by academics (see section 4.3.3.5.1).

Other aspects of evaluation systems caused concern, including the nature of the questionnaires themselves and of the students providing feedback (for example, Theme 2 responses to Q18 – see discussion in section 4.3.5). Perhaps as a consequence, then, in the current study some teachers saw the evaluation system as nothing more than a mandatory part of their institution's progression, appraisal and promotion system, and therefore engaged with evaluations in a minimal way; they administered evaluation questionnaires and reported results because of their institution's requirements. Even so, many teachers did report a focus on the evaluation system as a way of learning about the effectiveness of their teaching and courses, and engaged with evaluations in a deeper way, saying that they consider the students' responses to the surveys, discuss results with colleagues, and make changes to their courses in the light of the feedback (see Table 4:14 and section 4.3.5.2).

Alignment and misalignment – individual views, institutional claims, practices

There can be a mismatch between an individual's idea of evaluation and the institution's evaluation/appraisal system and processes. When there is a mismatch the resultant conflict may cause the individual to respond in a minimalist and even negative way, that is, be involved in the evaluation/appraisal system simply because it is mandated, but do nothing with the results of any evaluation/appraisal other than report them. On the other hand, the conflict may cause an individual to run an evaluation/appraisal, possibly because it is mandated, but focus on the results with a purely professional/development purpose in mind. Another response to the perceived gap between individual needs and institutional obligations is for the academic to use ongoing formative evaluation for developmental purposes, and participate in the formal appraisals as a ritual to demonstrate quality. Furthermore, when the focus of the institution appears to be on checking effectiveness and auditing quality, teachers feel that their professional capabilities are being undermined (Arthur, 2009; Burden, 2008) once again leading to the development of negative perceptions about evaluation (Moore & Kuol, 2005).

From the current study, it became clear that different staff engage with evaluation in different ways. Staff at OP whose jobs depend on good evaluations every year (see Table 2:5) were closely involved with evaluations in connection with teaching improvement, possibly because they are expected to report on how evaluations had contributed to the modifications/improvements they were planning for their courses and the changes they had made to their teaching. In the case of OP, there is direct alignment between the institution's mandate about student evaluation and the behaviours it expects its staff to demonstrate.

At WU, conducting appraisals and reporting the results are also part of mandated expectations for the annual Professional Goal Setting (PGS) process and in applications for promotion. However, the mandate is to report only appraisal scores within a larger Academic Staff Portfolio (ASP) on teaching, research and service. In the teaching section of the Portfolio there is no room for submission of any other data or description/explanation other than the appraisal scores (see section 2.3.1.2.4). Interestingly, of the WU staff who participated in the study, only a small number expressed negative views about their institution's use of the appraisal data (for example, three participants of the 20 who were interviewed voiced concerns – see section 4.3.3.5.2). If staff used their appraisal results for professional development, which many did (see Table 4:13), then it would depend on

whether the staff member also saw worth in using student evaluations/appraisals for that purpose. Even though at WU statements about evaluation/appraisal are included in a policy about evaluating teaching and courses in ways that point to the use of evaluation for professional development and improvement of teaching and curricula (see section 2.3.1.2.2), there are no explicit expectations that teaching staff *demonstrate* involvement in evaluation other than to report appraisal scores, which is part of a Human Resources policy related to the performance review. Thus, at WU there is a disconnection between the evaluation policy and expectation of behaviours.

There is even less direct alignment at OU, where there are no mandates about when or how often to run student evaluations, simply the expectation that teaching evaluation results and discussion of them are required to meet requirements for confirmation path or promotion. At OU, apart from general statements of advice/recommendation, formal statements about student evaluation appear only in Human Resources policies concerning confirmation path, promotion and annual performance review (see section 2.3.1.1.4).

The type of organisation – its mandate from the government, accompanied by how it is funded – determines the priorities and focus for the institution, which are then reflected in the mission, vision and goals statements (see section 2.2.1). As members of an institution, teachers are both contributors to, and subject to, the generation and continuation of the many policies, processes and practices governing all aspects of their work. Evaluation is part of that. If institutions are unclear about why they want their teaching staff to be involved in evaluation, and/or they do not communicate those reasons explicitly in ways that are understood, then confusion, lack of clarity and even mistrust and suspicion will result (Edström, 2008; Moore & Kuol, 2005). As a consequence, some teachers may respond blindly, following mandated instructions to run evaluations because they have to, or even avoid them altogether if they feel confident enough to be able to do so. Other teachers will endeavour to fit their institution's expectation into their own conceptualisation of evaluation, thus ensuring that student evaluation activities are not entirely meaningless. They will rationalise the meaningful evaluation-related activity with possibly conflicting or at least misaligned institutional demands.

5.2.2 Assertion 2

Assertion 2: Teachers' perceptions about the quality of student feedback including students' ability to make judgements about teaching and courses, influence how teachers view student evaluations and engage with evaluation/appraisal data.

This assertion is specifically about the relationship between section B of Figure 5.1 (*practical implementation*) and section C (*individual perceptions*): that teacher perception of evaluation processes, including views about the quality of the feedback that is collected from students, determines the nature of engagement with the data that are gathered.

Quality of the institution's student evaluation processes and survey instruments.

About three quarters of the participants in this study viewed student evaluations/appraisals as personally worthwhile (see Table 4:12). In addition, over half the respondents to Q19 of the questionnaire indicated that they found their institution's centralised student evaluation system effective at gathering useful/meaningful data for them (see Table 4:20). This is interesting, considering that much of the early literature on student evaluations focussed on questions about the validity and reliability of the student evaluation instruments as well as of student evaluators (Beran *et al.*, 2007; D'Appollonia & Abrami, 1997; McKeachie, 1990; Menges & Mathis, 1988; Millman, 1981; Smock & Crooks, 1973). Upon further examination of the data collected in the current study, the evidence showed that the claims of personal worth and effective centralised student evaluation systems were likely to be merely 'in principle' claims. Personal worth will mean different things to different people, of course, and recognition of highly worthwhile and high quality system efficiencies around evaluation may not necessarily mean much beyond acknowledgement of a well-run process.

Participants in this study saw evaluation data as being useful for a number of purposes, principally for helping them to know more about their students' learning experiences (see Table 4:13, Q2f), refine their teaching and their courses (see Table 4:13, Q2e), for professional development (see Table 4:13, Q2b), as well as its demonstration of teaching effectiveness for promotion (see Table 4:13, Q2c) or simply to meet requirements (see Table 4:13, Q2a). However, consideration of the evidence concerning the limitations of student feedback coupled with issues related to the evaluation systems and processes (see sections

4.3.5, 4.3.3.5, 4.3.1.2 and 5.2.3), shows that the generally positive views expressed in responses to Q2, Q4 and Q17 of the questionnaire needed to be examined more closely.

As illustrated in section 4.3.3.5, concerns about the evaluations process and/or the instrument emerged in themes in Section B of the questionnaire. The issues included design faults, the bluntness of the instruments and the difficulties with interpretation. In the interviews, the most commonly voiced concern was the notion that the timing of the formal student evaluations, which was generally at the end of the semester seriously limited their usefulness for teachers, students and the improvement of the teaching and learning experience (five at OP, three at OU and eight at WU). The unreliability or the limitations of the evaluations/appraisals instrument itself did not arise in the OP interviews, but eight interviewees at OU voiced criticisms of this nature, as did eight interviewees at WU. Issues included dissatisfaction with the questions, the number and representativeness of the students who answer the questionnaire and the need for more open-ended questions.

Quality of student feedback

As discussed in the literature review (Beran & Rokosh, 2009; Burden, 2008; McKeachie, 1990), evidence in the current study also indicated that the quality of feedback gathered through centralised evaluation/appraisal systems was viewed by teachers in a variety of ways, many seeing flaws including issues with the source of data, namely, the students (see discussion in section 4.3.3.5). One of the strands reported in the literature on evaluations over a long period of time is the perception of some academics that, for a range of apparent reasons, students' judgements are unreliable. Aleamoni (1981) synthesised some of the commonly stated reasons given for the view that students' judgements are flawed or unreliable. These include a lack of intellectual and personal maturity, a tendency to focus on likeability or immaturity and resistance to challenging material or assessments.

As the overview of common themes identified in Section B (Q9 to Q22) of the questionnaire outlined in section 4.3.3.5 illustrated, at all three institutions, there is a degree of uncertainty about the ability of students to evaluate. Between two per cent and 15 per cent of the comments made covered the general theme of doubting student ability to judge teaching. The interview data provided some further insights, as presented in section 4.3.5(a). The finding discussed in the latter part of section 4.3.5(a) – that just over half the interviewee group

across the three institutions were ambivalent in varying degrees about student capacity to judge – supports the sentiment expressed by Beran and Rokosh (2009) that academics were ‘neutral’ and ‘passive’ in their acceptance of student evaluations rather than enthusiastic. It should be noted, too, that while nearly half of the interviewees at the three institutions spoke positively about students’ capacity to judge, the group represents a small sample of volunteers who undertook the research questionnaire and volunteered to be interviewed (see Table 3:3).

Looking across the institutions, however, there are some noteworthy points about interviewees’ views of students’ ability to evaluate the quality of teaching and courses. The overall extent of doubt about students’ capacity to make judgements is significantly higher at OU (15) than at the WU (10). This may simply be the result of the sample of academics who agreed to be interviewed at each institution, but it is worth noting that the WU includes teaching explicitly in its Vision Statement, while the OU has the phrase “*A research-led University with an international reputation for excellence*” (see Table 2:4). As OU is an older and more traditional university with a research emphasis, it is perhaps not surprising to find that academics express stronger reservations about the role students are able to play in the educational processes.

Another common thread that is reported in the literature is the view held by academics that students do not have the appropriate tools and knowledge to make judgements. One often-reported perception is that students are liable to confuse personally attractive dimensions, such as charm, entertainment and likeability with genuine educational worth, although the nature of the latter is generally undefined. This research study indicated that this is a suspicion that some tertiary teachers still hold. The numbers of interviewees at OP who referred to this point in some way was not high (four), but the suspicion that students’ evaluative feedback is shaped by popularity or likeability was noticeably high at OU. Nine interviewees voiced this concern at OU, which is in keeping with the strong expression of doubt about students’ capacity to judge that emerged in the OU findings. At WU, respondents who questioned students’ competency to judge did not emphasise the popularity factor – only one pointed to this explicitly – but mentioned a range of other limiting factors. There is no obvious explanation for the preponderance of popularity concerns amongst the interviewee group at OU as opposed to the other two institutions, but it is interesting to note that this was

also accompanied by a significantly higher degree of suspicion about how academics might use the evaluation instrument for their own ends (see discussion in section 4.3.3.5.2).

Other questions about students' evaluative competency that support those reported in the literature occurred in varying degrees in both the interview data and the questionnaire comments (see section 4.3.3.5). The interview data, in particular, highlighted different rationales for the perceptions of students at the three institutions. For example, some interviewees at OP talked about students lacking the necessary tools to judge and cited factors that they saw as specific to their context. One interviewee suggested that students in the pre-degree programmes would not be able to articulate their views on these matters, while another argued that, for some young students, judgement on the teacher would be deemed inappropriate. Further students' shortcomings that were cited by OU interviewees included immaturity, an inability to grasp course goals, failure to cope with challenging material and the absence of a long-term perspective. All of these reservations also surfaced in the comments of WU interviewees, as well as views about some other student limitations. In the WU group, two interviewees additionally raised the impact of fee payment on students' evaluations, the superficiality of students' feedback and the notion that some students used evaluations to express personal dislike of a lecturer.

These findings, and the repetition of certain themes, suggest that there are narratives around student evaluations that have become deeply embedded in the minds of tertiary teachers although these are not necessarily supported by evidence. Interestingly, these narratives seem more prevalent in the universities where perhaps the large class traditional lecture format promotes notions of students as a collective entity with certain assumed characteristics.

Concerns about the quality of student evaluation data similar to those identified by Aleamoni (1981) were shown to exist among the participants in the current study. Students were seen as not being in a position to be able to make judgements about teaching, with particular note being made of their lack of experience and knowledge of what is good in terms of teaching but also in terms of content (*e.g.*, see Q18, sub-theme 2d and the discussion in section 4.3.3.5). Participants in the current study also expressed the concern that students were easily swayed by easy courses and likeable teachers; essentially seeing the whole evaluation system as a 'popularity contest'. This view was expressed also by a proportion of all interviewee

groups from across the three institutions, but particularly strong with those from OU (see section 4.3.1.2).

Students, survey instruments, curriculum development

A criticism about the nature of student evaluation/appraisal data was that they were of limited use for helping improve or enhance teaching and courses, a criticism that seemed to be levelled at both the instruments and the source of the data, the students (see discussion in section 4.3.3.5). This issue was also highlighted in the study by Burden (2008). Amongst other things, the survey instruments were seen as being too blunt, not statistically reliable and the lack of comparative data meant that teachers were not able to monitor change, or see how they ‘measured up’ against other teachers. The concern with students’ limited ability to contribute worthwhile and meaningful feedback about teaching and their learning has already been discussed.

Interestingly, the student evaluation/appraisal surveys run at the three institutions participating in this study, as with many of the student evaluation surveys that are run in tertiary institutions, are designed to collect data about students’ experiences. As discussed in section 4.3.3.5, on the one hand, some teachers in the study were critical of survey instruments because they were too blunt and did not collect the breadth of data needed to inform teaching and learning. Some teachers criticised students as not being able to contribute reliable feedback on teaching because their views can be influenced by other assumedly irrelevant factors such as emotional response. It is understandable then, why current student evaluation surveys are unsatisfactory for teachers who hold these views. The surveys are designed to collect information about student experiences, yet these teachers expect, and perhaps assume, that they collect more than that.

One implication of these criticisms could be that for those who view students as not being capable of providing reliable feedback on teaching, evaluation of teaching should not involve students at all. Another implication could be that student evaluation surveys should be designed to gather feedback from students, as well as other sources, to seek a more rounded view of teaching. Some participants did point out that a number of different methods of data gathering should be included in any evaluation system (see discussion in section 5.2.5).

In summary, Table 5:2 lists what seem to be the main concerns related to students and student evaluation survey instruments in the form of assumptions or expectations and the possible criticisms of the current systems.

Table 5:2 Assumptions and Possible Criticisms about Students as Sources of Data, Student Evaluation Systems, Processes and Instruments

Assumptions/expectations	Possible criticisms of current systems
Surveys evaluate the whole teaching/learning environment	There are flaws in the instruments because they do not evaluate the whole teaching/learning environment
Surveys do not evaluate the whole teaching/learning environment. They only evaluate students' learning experience	Surveys are unhelpful for developing and enhancing teaching
Surveys are not statistically reliable. No comparative data are produced	Surveys are of little use for gauging progress or change and cannot support teaching and learning development
Students not able to provide feedback on teaching and learning because of limited experience/knowledge	Evaluation processes are flawed and the system is not helpful for developing and enhancing teaching. Student feedback is not valued
Students are biased and are unreliable as sources of data	Surveys and students are of little value. Data are untrustworthy and cannot help with teaching and learning development

Policy versus implementation

The question becomes, of what is this circularity of argument a symptom? Perhaps one cause is that there is a variety of expectations held by institutions and by individual teachers about centralised evaluation systems and the place of evaluation in teaching. It could also be related to the disconnection between the view of student evaluation that institutions portray (that evaluation is principally about development) and the view of evaluation that the institution enacts (that the results of student evaluation surveys are the *primary* way to demonstrate effective teaching).

Institutional statements/policies indicate that student evaluation surveys only gather one kind of feedback, namely, feedback from students on their learning experience, and therefore there is a need to supplement evaluation survey data with other data. At OU for example, a reflective discussion about the evaluation results in the light of the academic's context is a

requirement in the *Otago Teaching Profile* (see section 2.3.1.1.4). At WU there is guidance about the including of a variety of forms of data in reflection on teaching (see http://www.waikato.ac.nz/tdu/pdf/booklets/15_Appraisals.pdf). At OP, where there is greater alignment between policy and practice than appears to be the case at the two universities, it seems, its *Student Surveys and Course Evaluations* policy states explicitly that student evaluations contribute just one set of data (see http://www.otagopolytechnic.ac.nz/fileadmin/DepartmentalResources/Marketing/Policies/Academic/AP0700.06_Student_Surveys_and_Course_Evaluations.pdf). However, implementation of these policies and statements do not portray the same view of evaluation in practice. At OU, for example, the only compulsory inclusion of student evaluation data in the *Otago Teaching Profile* is student evaluation of *teaching* data, and course evaluation is optional. At WU, to demonstrate effective teaching, only the results of the teaching appraisals are required to be submitted. There is no expectation that teachers demonstrate their engagement in reflection and professional development using appraisal data. At OP, the *Student Surveys and Course Evaluations* policy gives primacy to teaching evaluation data and major decisions about intervention are based on teaching evaluation scores. Having said this, however, teachers at OP are expected to show how they use evaluations to refine and enhance their courses.

The variety of perspectives about the worth and quality of student feedback, including views about students' capacity to provide meaningful evaluation found in the data gathered in the current study, could very well be related to the degree of match and mismatch among: institutional expectations about student evaluation as claimed in the three institutions' policies and guidelines; their implementation through process; the nature of the survey instruments; and teachers' interpretation of the whole system in light of their own views, beliefs and practices about evaluation.

5.2.3 Assertion 3

Assertion 3: The use of the same instrument for quality purposes and to inform teaching influences teachers' views of evaluation and their evaluation-related behaviours.

This assertion concerns section B of Figure 5.1 (*practical implementation*) in terms of how the use of one instrument to gather student feedback impacts upon *engagement* with evaluation/appraisal (section D in Figure 5.1).

One instrument, two purposes?

Stronge (2006) describes evaluation as a way to improve performance while simultaneously providing people opportunity to demonstrate accountability for what they do. In principle, this may seem like a logical approach to take with student evaluation in tertiary institutions. Indeed, as Bowden and Marton (1998) suggest, “if improvement is addressed properly, evidence for accountability will be developed automatically” (p. 228). In many tertiary institutions, however, Stronge’s ‘simultaneously’ means the use of a single instrument (even though there is often a degree of in-built flexibility, such as with the instruments used at three institutions participated in this study – see Table 2:5) to achieve two purposes.

Edström (2008) highlights the tension or conflict that can arise when the same instrument is used for auditing purposes as well as for the purpose of development and reflection. Indeed, a strong sense of this tension formed part of the rationale for this study. In essence a disconnection can occur between quality as ‘accounting’ and quality as ‘accountability’. This conflict can weaken the potential of student evaluation/appraisal instruments and processes to contribute to the improvement of teaching and courses. In addition, as Arthur (2009) points out, the very existence of a centralised evaluation system that has an audit focus (or that is perceived to have an audit focus) undermines the professional autonomy of teachers and thus causes them to be less willing to use student evaluations to enhance their teaching and courses, a point also highlighted by Burden (2008).

Institutions rely on policy to state and communicate intentions and purposes behind action. In larger organisations especially, teachers tend to rely on one another to find out about the practical implementation of policies. Experiences and interpretations are shared, sometimes generating myths and stories about the institution’s purposes and intentions, that may not match the intentions as stated in policies. So, while a policy statement might seem in line with currently accepted best practice, it is the staff interpretation of the policy that is the reality of how the policy is implemented. As Edström (2008) reminds us, “It doesn’t matter

much what the institution's *intended* purpose is. What is important is what the individual teachers *perceive* to be the purpose" (p. 100) (emphasis in original).

As has already been discussed in section 5.2.2, the three institutions have evaluation policies or statements in place and for all three there seems to be some disconnection between those policies and statements and their interpretation and implementation (by institutions in their creation and carrying out of processes, and by teaching staff engaging in those processes). However, while there may be logic behind using one evaluation instrument for two purposes, and researchers would argue the possibilities, especially if there is a focus on ultimate improvement or enhancement (*e.g.*, Ramsden, 1992), putting mechanisms in place and having them understood in a shared way by the institution and individual teachers alike is not straightforward – “The same data seldom serves both purposes well” (Patton, 1997, p. 78).

Two systems, two purposes?

Separating the two processes may be a way forward. For staff who feel particularly vulnerable (because they are on confirmation path, *e.g.*, at OU, or are subject to annual Professional Goal Setting review, *e.g.*, at WU, or their position is reliant on the outcomes of student evaluation processes, *e.g.*, at OP), this separation will allow them to engage in evaluation in a way that makes sense from a curriculum/teacher development point of view that is, arguably, the original purpose of evaluation as it relates to teaching. Evaluation is an integral part of curriculum implementation. Evaluation informs those who know the context of the effectiveness of the curriculum. If there is the ever-present risk that judgements are being made by those with power over an individual teacher's future, then it is likely that the teacher will not engage in the evaluation process in the best way he or she can. That is not to say that the process should be an individual and a hidden one, but data from evaluation processes are essentially for the immediate stakeholders in the context. Those who know the context intimately understand where the data have been generated, and they and colleagues are in the best position to respond to emerging needs or issues. Separating the two processes would allow teachers to develop key skills of effective teachers, namely, using Eisner's (1985) terms, their *connoisseurship* and *critical ability*:

If connoisseurship is the art of appreciation, criticism is the art of disclosure.

Criticism...has at its end the re-education of perception... The task of the critic is to

help us to see...Thus...connoisseurship provides criticism with its subject matter. Connoisseurship is private, but criticism is public. Connoisseurs simply need to appreciate what they encounter. Critics, however, must render these qualities vivid by the artful use of critical disclosure (pp. 92–93).

Participants in the current study pointed out that student evaluation data do not provide enough of a picture of teaching and courses (see discussion in section 4.3.3.5) and that they regularly implement informal and more regular evaluation strategies to fill gaps and to provide more meaningful information. Essentially, these teachers recognised that the evaluation process operated centrally by their institution is a ‘pared back’ version of evaluation, as centralised systems only capture a snapshot view from one group of stakeholders involved in the learning environment, namely, the students.

Separating the two purposes – and making the purposes clear to all (Edström, 2008) – may provide the basis for reducing confusion, increasing understanding, and a clearer foundation for representing the concept of student evaluation and evaluation processes more broadly. This separation will enable staff to make better use of centralised evaluation systems because they will be more sure about the rationale behind the data gathering, analysis and outcomes. They would also be more sure about when the process involved scrutiny by the institution about performance and effectiveness and when they could concentrate their efforts on teaching, course and self-development.

One integrated system, two integrated purposes?

Rather than separating the two processes, perhaps leaving the system as is, while clarifying purposes, expectations and how the data can have meaning for all involved may be another approach that is practicable. This would encompass breaking down the disconnection between quality as ‘accounting’ and quality as ‘accountability’. In the current study, teachers did acknowledge the various uses of student evaluation/appraisal data and said they used the data to achieve a variety of purposes that encompass both determining quality and for development (*e.g.*, see Table 4:13). There were staff who seemed able to accept the varying purposes and rationalise both.

However, “collecting data is not the same as improving or judging teaching” (Ramsden, 1992, p. 232) and using the data collected via student evaluations is not necessarily straightforward, whether using data for professional development/improvement or to measure effectiveness and performance. While it is acknowledged that using student evaluation/appraisal data for either of the two purposes can cause issues, Smith (2008) suggests that they can indeed work alongside each other, if student evaluation/appraisal data are viewed in a more systematic way by institutions, and individuals embed their use of data within their reflective practice. He suggests an arrangement in which there is systematic interpretive guidance about student evaluation/appraisal data provided for teachers, as well as opportunities for staff to engage in communities of learning about their teaching and their courses in the light of student evaluation processes and outcomes. As discussed already in the literature review, Smith’s (2008) five-phase programme, drawing on four types of evaluative data (not just student evaluation/appraisal of teaching/course data) linking staff development and evaluation, is one way to meet both institution and individual teacher needs. While Smith’s ideas are not necessarily new (*e.g.*, see Ramsden & Dodds, 1989), it is the systematic and integrative nature of Smith’s programme that is different. In the case of the three institutions that participated in this study, following Smith’s (2008) advice would result in the continuation of one instrument being used for evaluation/appraisal, complemented by other forms of data gathering and embedded within a systematic framework of consultation and support. The institutions themselves would ‘tap into’ the results of student evaluations. They would need to provide resourcing for individual teacher and teaching community-integrated reflection and interpretation of scores, and build up a meaningful trail of student evaluation/appraisal results to provide some representation of the overall status of teaching, as well as inform its future development.

This study indicated that current systems of student evaluation using one instrument have produced tension and limited the developmental potential of formal evaluations. Changing the systems to separate the processes of development and audit may be a way forward. On the other hand, streamlining the current systems to ensure there is a match between intents, purposes and practices supported by resources may also be a way to reduce confusion and possible alienation felt by teachers.

5.2.4 Assertion 4

Assertion 4: Teachers' beliefs about teaching and learning and how they view themselves and their role in their institution will influence their perceptions of, and engagement with, student evaluation/appraisal.

This assertion is specifically about links between section C of Figure 5.1 (*individual perceptions*) and section A (*institutional contexts*), and the effect of those links on teachers' engagement with evaluation/appraisal (section D).

Teaching beliefs, teacher roles

As discussed in the literature review, views held about teaching and learning may underpin and therefore influence and determine beliefs and practices related to evaluation. While only one study was cited that had investigated this link specifically (Hendry *et al.*, 2007), it stands to reason that based upon the solid research base around conceptions of teaching and their link with teaching practice and student learning (*e.g.*, Biggs & Tang, 2007; Ho, Watkins & Kelly, 2001; Pratt, 1992; Ramsden, 1992; Trigwell & Prosser, 1996a; 1996b), conceptions of teaching would also influence evaluation practices, especially as evaluation is an integral part of teaching activity.

Logically, such views have an impact on the way a teacher sees his/her role and the sense he/she makes of the institutional requirements and expectations, policies, systems and processes around evaluation. Beliefs expressed about teaching and learning varied across the data, from student-centred to teacher-centred. Variations in beliefs were identifiable across all institutions (see discussion in section 4.3.1). In a student-centred/focussed view, students are seen in partnership with the teacher, rather than separate from the teaching-learning environment. As a consequence, teachers tend to value student feedback and see the enhancement of the learning environment as a collaborative venture. In other words, when teachers see learning and teaching as involving processes that are student-focussed, there is a greater chance that the teacher will also see evaluation as involving students in an integrated way. This view of the nature of teaching as being a collaborative endeavour often also includes colleagues as well as students. Some participants in this study indicated that sharing evaluation/appraisal results with colleagues was a natural activity (see Q4e, Table 4:14).

Course and paper development also featured strongly as a reason (see Q2e, Table 4:13) for running student evaluations/appraisals. For some teachers, the centralised student evaluation system provides one means of gathering feedback from students, not the only means. Other strategies were described that were implemented more often and less formally, to allow teachers to gain evaluative feedback on aspects of the learning environment they thought were directly relevant to their context, and in a timely fashion, so there was opportunity to discuss the feedback with the students who had contributed. This theme appeared in all the comments questions from Section B of the questionnaire (Q10, 12, 14, 16, 18, 20, 21 and 22) and examples are given in the additional details section of each comment data analysis table (see Appendix 6). There were also comments made in the interviews about the importance and place of gathering evaluation data from a variety of sources (see further discussion in section 5.2.5). These comments highlighted the need for other evaluation strategies to complement formal student evaluation instruments, such as, peer and expert review, class representatives, informal dialogue, informal surveys, quizzes and focus groups (see discussion in section 4.3.5). Teachers could monitor ongoing developments and respond in a practical ways throughout the rest of the course, rather than wait for the end of the semester.

Thus, in this study, there emerged evidence of teachers who expressed student-centred views about teaching and learning as well as teachers who expressed more teacher-centred views (see discussion in section 4.3.1.2). The findings from the three institutions indicate interesting commonalities and differences in relation to the three research questions and the interplay between the three facets of views of teaching and learning, the capacity of students to provide worthwhile feedback on teaching and courses and how teachers engage with evaluations. In terms of the articulation of teaching and learning beliefs, the most striking difference was between the OP and the two universities. The interviewees at OP generally provided detailed reflections on their teaching conceptions and were also more inclined to refer to educational theorists. Significantly, there was a strong focus on students in their accounts of the teaching and learning process and on preparing students for participation in the work force or the community after tertiary study. Alongside the focus on learner outcomes, a number of interviewees referred to the importance of practical learning. Research concerns did not feature in the discussion of teaching practices, but many interviewees referred to transitioning

from the role of practitioner to a teacher role. In addition, OP interviewees described a much higher degree of engagement with student feedback.

Interviewees from the two universities did not demonstrate the same focus on students and their long-term outcomes as the OP interviewees. At both universities, interviewees who focussed on students and articulated well-developed goals for their students beyond the university were more exceptional. At the universities, there is still a common tendency to talk about teaching and learning in terms of teacher behaviours and/or the handing over of content. Additionally, some interviewees still see themselves primarily in terms of their discipline research.

The difference in emphasis between OP and the universities in relation to students and content when teaching and learning beliefs are discussed is not surprising in terms of the historical antecedents of the different tertiary environments, as well as their current goals and student populations. Polytechnics have always been geared towards preparing for the workforce and they tend to put a strong emphasis on teaching. Many of their teaching staff members have been, or are, practitioners in their field. In New Zealand, even today, most university academics are appointed without a teaching qualification, but on the basis of their discipline research. This difference is still apparent in the way that more polytechnic teachers conceptualise their role in terms of outcomes for the students. However, a student-focussed approach appears to be becoming more common than previously in the university context. At OU, eight interviewees mentioned the students in their teaching conceptions, although some only did so in peripheral ways. Likewise, at WU, eight interviewees explained their teacher understanding in terms of a primary focus on students, but again for some this was a considered and developed position while for others the reference to students in the learning process was more cursory. Nevertheless, in both universities there were clear signs of academics moving away from the content-delivery model of teaching and learning.

Teaching beliefs, faith in student judgement

Hendry *et al.*, (2007) suggested that there is a link between teachers' perceptions of teaching and learning and their responsiveness to student feedback. Their study indicated that teachers with a student-focussed approach were more likely to be responsive to student feedback than those who adopted an information-transmission stance. Their study investigated the

possibility of a correlation between teaching beliefs and attitudes to, and practices around, student feedback from a number of angles.

With regard to a link between teaching and learning belief statements and views of students' competency to judge (see also section 5.2.2), the findings in the current study showed a surprising pattern. In spite of the very emphatic focus on student learning needs and outcomes at OP, 13 out of 20 interviewees expressed some degree of ambivalence about students' capacity to judge the quality of the teaching and their learning experiences. By contrast, while the findings from the universities showed a much smaller number of interviewees who were focussed on students and their learning outcomes, there was a close correlation between those with student-focussed approaches and faith expressed in students' judgements. At OU, all three respondents who were very positive about students' competence to make judgements, prioritised students in their articulation of their teaching and learning beliefs. This was also true for the two OU respondents who were mainly positive about students' competence to judge, and those who were primarily positive with some reservations. Likewise, at WU, the correlation between developed, student-focussed beliefs and faith in students' capacity to make judgements was also evident. Five of the ten interviewees who articulated conceptions of teaching with a strong student focus were strongly supportive of students' evaluation competence.

The difference between the universities and the polytechnic in the link between beliefs and attitudes about students' judgements can be addressed only speculatively. It is possible in the polytechnic environment, which has always had a strong student and teaching focus, that reflective student-centred thinking about teaching is the norm (as evidenced by the preponderance of belief statements of this ilk). However, in the more research-focussed university environment, well-developed and student-focussed teaching and learning beliefs are still fairly exceptional and may reflect a deliberate and purposeful engagement with pedagogy on the part of some individuals. In this context, it is not surprising that student-focussed beliefs were often matched by faith in students' evaluative ability. The people who expressed these beliefs in the university have had to challenge cultural norms in the institution unlike those in the polytechnic and may therefore have markedly different perceptions about students and their role in the teaching and learning process.

Teaching beliefs, teacher roles, institutional expectations

There are many contextual factors that may help to explain these indicators of a gradual shift from a content-defined notion of teaching and learning to a student outcome approach.

Contextual factors include the graduate attributes/outcomes and quality agendas and the associated demands of stakeholders to see identifiable outcomes of tertiary study.

Such contextual factors have led to institutional requirements about student evaluations/appraisals being mandated, and while this was one of the reasons for running student evaluations acknowledged by questionnaire respondents (85 per cent of respondents to Q2a of the questionnaire – see Table 4:16), it was not the primary reason. Data showed that getting feedback on students' learning experiences (Q2f, 93 per cent) and helping with paper refinements (Q2e, 89 per cent) attracted a higher percentage of respondents. As was discussed in section 4.3.1, these high 'yes' responses indicate that teachers claim interest in activities that are more aligned with professional development and fine tuning their teaching and courses, which, arguably, may demonstrate some sense of student-centred views. In contrast, however, providing feedback to students, along with reporting to internal and external bodies, was the purpose for which respondents said they least used student evaluations/appraisals (Q2g, 44 per cent and Q2h, 46 per cent respectively). Not involving students may indicate a more teacher-centred view than a student-centred one, or simply an unsophisticated view of the possibilities for using student feedback.

For those teachers who were at particular points in their careers, there were some indications that moving through confirmation path (OU) or having to demonstrate effectiveness on a regular basis (all three institutions) or applying for promotion (WU, OU) or salary review (OP) meant that there was a heightened sensitivity to institutional expectations (*e.g.*, see section 4.3.3.2). For these teachers, balancing institutional expectations with personal beliefs about teaching and learning and evaluation was not always straightforward. The discussion of teaching beliefs in section 4.3.1.2 highlights this. For some, institutional expectations sat alongside personal beliefs about evaluation and its use as an integral part of curriculum implementation and the associated 'natural' processes of reflection and development. Indeed, the two purposes (development and quality assurance) were seen as complementary or at least not in opposition to each other. For others, the two purposes were seen as clashing, with teachers tending to see evaluations as being for one purpose more than for the other purpose. Some teachers who saw evaluations as being principally for development were also not

interested in career advancement (they had already reached their goal in that regard) or were keen to learn about their teaching and courses because it was the development of that aspect that their institution expected its staff to concentrate on, as with OP.

Teachers may feel able to use student feedback, whether positive or negative, because they view teaching as a student-centred activity and therefore see value in student feedback. On the other hand, they may be in a position of not knowing how to interpret feedback, or even avoiding it if negative; perhaps because their basic views about teaching are not student-centred and they therefore do not see value in student feedback. Some comments made in the interviews indicated this (see section 4.3.1.2).

The isolation of institutional evaluation systems from the heart of the teaching and learning process, often seen as an individual or personal endeavour, also influenced teachers' views. Difficulties in seeing how evaluation formed part of a bigger system of feedback beyond the immediacy of the closed teaching context added to that perspective. As well, limited embedded institutional support for learning from feedback served to exacerbate teachers' passive or negative feelings about evaluation and even encourage feelings of personal alienation from the wider institutional context (see discussion in section 4.3.5). This was demonstrated to some degree by the data showing that more OP people were sensitive to, and more than probably personally engaged with, the issues of needing to report evaluations to external and internal bodies (see Table 4:13). Many of the teachers would be involved in creating those reports and would therefore have a higher awareness of institutional needs to have access to student evaluation data. In the larger organisation, such as the two universities in this study, reporting is often not the role of the teacher and so teachers do not necessarily know what reporting entails.

5.2.5 Assertion 5

Assertion 5: Teachers tend to view student evaluation/appraisal as an isolated and individual activity that informs them about their courses and provides data to demonstrate their teaching effectiveness to their institutions.

This assertion focuses particularly on section C of Figure 5.1 (*individual perceptions*) highlighting teacher behaviours or the way they engage in evaluation/appraisal (section D). This assertion also points to how the limited notions of 'evaluation' that seem to underpin

student evaluation instruments and processes, influence the way teachers view and understand evaluation and its role in teaching, learning and curriculum development.

Evaluation – for self-interest

While there are studies which draw attention to the negative views teachers have of student evaluations/appraisals (*e.g.*, Aleamoni, 1981; Arthur, 2009; Moore & Kuol, 2005), as well as those that highlight the more positive worth teachers place on them (*e.g.*, Nasser & Fresko, 2002; Penny & Coe, 2004; Schmelkin *et al.*, 1997), the current study raised interesting facets of the ‘grey area’ between the negative and the more positive views. Participants in this study held generally positive views about the worth of evaluations/appraisals. However, as discussed in section 4.2.1 and above, the worth of evaluations was not necessarily linked to using them to enhance teaching, learning and courses.

The participants in this study saw evaluation variously as:

- *a way to meet requirements.* This view was accompanied by a focus on reporting to others, implying a need to be involved in evaluations/appraisals because others need to know the results. The extent of engagement seemed to be one way of demonstrating to others that regular evaluation takes place and that results are acceptable
- *a way to promote oneself.* This view is held by those whose aim is to report that evaluations/appraisal have occurred in line with institutional expectations, the principal purpose being self-promotion. By achieving good results, it can be demonstrated that institutional demands have been met and/or that results deserve recognition, through promotion or positive review. The focus of engagement in evaluations/appraisals is, essentially, one of self-interest
- *a way to get to know what is going on.* This view focuses on teachers knowing how teaching and courses are progressing in the eyes of the students, simply for the sake of knowing, not to make changes. There is sense in this view that student evaluations are run to satisfy institutional (and possibly student) expectations and as long as, in general, there are no issues raised, feedback is not looked at closely and not taken on board in any serious way. The ultimate intent is not to learn/build/develop teaching, but simply to know

- *a way to determine the effectiveness* of the course and the teaching with a view to building, developing, creating and changing teaching, learning and curricula. For those holding this view, engagement involves both teacher self-development (short-term and long-term) and course development, including within that, the development of better learning environments for students. Engagement will often involve students and colleagues as well.

As was discussed in section 4.2.1, many teachers in this study saw value in student evaluation/appraisal data if the data were able to be used for the purposes that they thought were important, usually to assist them with their teaching and their courses or to get through the system, in terms of reporting, or for reviews, confirmation and promotion. Even so, though a high value seemed to be placed on learning about student experiences and improving courses, relatively few participants said a purpose of student evaluations/appraisals was to feed back to students (see Q2g in Table 4:13), that they share or communicate evaluation/appraisal results with students (see Q4c in Table 4:14 and Q6 in Table 4:24) or discuss results with colleagues (see Q4f in Table 4:14). How evaluations serve the needs of students thus did not appear to be the focus of the perceptions of staff, as observed during this study.

Similar to the argument of Beran and Rokosh (2009), it would seem perhaps, that while participants said they found evaluations/appraisals personally worthwhile, further evidence showed that it does not necessarily follow that they use the evaluations to enhance their teaching. Centralised student evaluation systems seem to have originated in an era when particular emphasis was being placed on encouraging tertiary teachers to think about the importance and place of the student experience/perspective, that is, to take a more student-centred view of the educational endeavour. The introduction of formal student evaluation questionnaires was one way to achieve that change. At the same time, institutions saw an opportunity to use the data to help determine teaching quality. Indeed, research has demonstrated that the quality of teaching and learning can be indicated by data gathered through student feedback questionnaires (*e.g.*, Kember *et al.*, 2002; Marsh, 1987), offering convincing evidence that not only could an evaluation system encourage teachers to place more value on the student perspective, it could also inform the institution of the effectiveness of teaching staff.

However, measuring quality and using the process to improve quality are two different things. While the connection between student evaluation questionnaires and improved teaching may be possible (*e.g.*, Kember *et al.*, 2002), it does remain tentative (Beran & Rokosh, 2009; Nasser & Fresko, 2002; Schmelkin *et al.*, 1997), because of the many purposes teachers have in mind when they run student evaluations. For many teachers in this study, if ‘getting through the system’ – by using data to review teaching and courses, to demonstrate effectiveness, or both – is the primary purpose behind engagement in evaluation, then the focus will be on that purpose rather than on professional development or enhancement of teaching. As a consequence, such teachers may have a tendency to perform ritualistic evaluation behaviours, rather than educationally meaningful ones (Ory & Ryan, 2001). Together with a teacher-centred view of teaching, learning and evaluation, this minimalist engagement in evaluation and review results in students and colleagues being less likely to be included. In addition, the way centralised student evaluation systems represent evaluation and its place in teaching and curriculum can portray and promote this impoverished view. For example, as Edström (2008) observed, while student evaluation systems and questionnaires claim to support the development of courses, they resemble auditing processes and instruments.

Evaluation – an isolated and emotional activity

As has been noted, a number of contemporary studies on student evaluations have challenged the notion that academics are opposed to student evaluations (Beran & Rokosh, 2009). At the same time, commentators observe that academics’ acquiescence to student evaluations often does not convert into using them to improve teaching (Arthur, 2009; Beran and Rokosh, 2009; Penny & Coe, 2004; Smith, 2008). The gap between acceptance and engagement appears to have many causes, and it is speculated that the presence of strong emotion around receiving appraisals may be a factor (Arthur, 2009; Moore & Kuol, 2005) In the course of the interviews, interviewers tried to probe this emotional response by asking a direct question, “*when you receive the results of appraisals, how do you feel?*”. In all three institutions, the responses to this question were less than either the literature or hearsay reports would lead one to expect.

However, when interviewees acknowledged the presence of emotions, the comments often suggested emotional rawness (see discussion in section 4.3.1.2). At OP, six respondents

freely articulated their sensitivity to criticism, while at OU the language used by a number of respondents, whether positive or negative, indicated the emotional contours of the evaluations/appraisals terrain (use of language like 'excited', 'vulnerable', 'crushed', and 'fear'). Again at WU, many respondents used language that indicated emotion, whether positive or negative, with some using language that captured their emotions and others more directly talking about the affective impact of evaluations. Only three respondents at WU explicitly discussed the negative impact of evaluation on their personal and emotional wellbeing. The presence of emotionally-tinged language in the responses of interviewees at all three institutions supports the views of Moore and Kuol (2005) and Arthur (2009) that emotions play a part in receptivity to evaluations. At the same time, the notion of evaluations as something personally distressing was explicitly stated by 13 out of the 60 interviewees (six at OP, four at OU and three at WU). Thus, while there was evidence of heightened emotion in the language choices of interviewees, fewer than anticipated discussed negative emotional impact directly. However, it is arguable that interviewees may have felt uncomfortable voicing their negative emotions directly or the formulation of the question may have been unclear. On the basis of the findings, this study supports the idea of a strong emotional tenor on the topic of evaluations/appraisals with a small number of interviewees alluding to it directly. These findings suggest that there is still research to be done to obtain more conclusive evidence on the impact of emotions on academics' use of student evaluations. Although no conclusive evidence was found, these reactions to receiving feedback, especially when negative, could be an additional reason why evaluation seems to be an individual and private activity for the teachers who participated (Arthur, 2009; Moore & Kuol, 2005).

Traditionally, teaching has been seen as a solitary activity (Laufgraben & Tompkins, 2004; Lucas, 1996) and even though calls for a change to a view of teaching as a collaborative and shared enterprise are not new, the individuality of the teacher and teaching persists. Added to that, in tertiary organisations, it is relatively easy for a teacher to maintain a solitary approach: academics are independent people on the whole and principles such as 'academic freedom' can work against collaborative efforts. The size of the institution can also contribute with many teachers seeing themselves as part of a discipline, department or school rather than an institution (Becher & Trowler, 2001). Some teachers in the study stated that they were not aware how their institution used student evaluation data. In the case of OP, with its 'flatter'

structure and smaller size than the two universities, it appeared that teachers were more aware of institutional use of student evaluation data, one reason being, that they were more closely involved in reviewing student feedback with the aim of reporting to external bodies (see section 4.3.3.3).

Evaluation – involving collaboration and more than one form of data

The literature suggested that many academics believe that student evaluations should be complemented by other forms of evaluation such as peer and self-review (Ballantyne *et al.*, 2000; Braskamp & Ory, 1994; Penny & Coe, 2004). In section 5.2.3, Smith's (2008) "four quadrant" approach was discussed as a way of mapping the different data sources that can be drawn upon to feed into evaluation. The four data sources are: student experience; self-review; student learning and peer review (p. 527). Only one of those sources, student experience, concerns data gathered through student evaluation systems, but together, the four different sources form a picture of teaching and courses that is richer and more accurate than the picture that only one form of data can provide. Conceptualising evaluation in this way presents evaluation as being more complex and integrative than the view that is often portrayed when the focus of evaluation is perceived to be on student evaluations/appraisals alone. It also highlights the importance of teaching as a collaborative and shared venture, rather than an individual one, encompassing both students and community groups of teachers.

The importance of obtaining complementary evaluative feedback from other sources was a theme that appeared in both the questionnaire and interview data. However, in relation to the questionnaire comment data in response to Section B questions, while this theme appeared consistently, it was not by a large percentage of staff (percentage responses varied between two per cent and nine per cent). Respondents also raised the idea that formal summative evaluations should be accompanied by getting other kinds of feedback on students' learning progress. A number of interviewees at all three institutions made a connection between their perceptions and use of summative evaluations and their reliance on other forms of ongoing feedback. This was particularly striking in the comments of interviewees from OP who talked about using small classes and the practical learning environments to keep in touch with their students' learning progress (five) as well as those who deliberately set up other methods of

formal evaluation (nine). This high percentage of interviewees who spoke about ongoing engagement with their students and their learning progress reflects a culture and a context in which teachers appear to work alongside their students to a far greater extent than at the universities. The feedback about proximity to students' learning is also in keeping with the predominance of student-focussed views in the teaching beliefs of OP interviewees. The implications of this close, ongoing engagement with students' learning are interesting. There was a sense that the focus on summative evaluations is less intense because of prior knowledge of the students' learning. For some of the participants, highlighted through the interviews, this meant that the evaluations are mainly used to validate or confirm what is already known by teachers. In other instances, lecturers expressed little interest in the formal summative evaluations, because they felt that they were far less valuable than the formative evaluations they had conducted or the inter-personal connections they had built with their students.

The reference to other forms of student feedback was also evident at the universities, although less strikingly than at OP. At OU, six interviewees talked about the need for formal evaluations to be complemented by other forms of evaluation. Some of these participants went on to discuss the formal evaluations as one element in a total feedback and evaluation picture. Others felt that the real learning came from formative evaluations, so that they felt comfortable about ignoring the professional development opportunities that may be available in summative evaluations. Four WU interviewees spoke about formative evaluations. While the idea of ongoing evaluation is definitely present in the statements made by both university interview groups, there was not the same suggestion of a common learning endeavour with the students as in the interviews held with OP teachers. As OP interviewees mentioned, small classes and practical learning contexts are conducive to greater ongoing engagement with students.

As already stated, the ways in which evaluations serve the needs of students is not the focus of the perceptions of staff, as observed during this study. This was indicated by the relatively low percentage of respondents saying that they always or frequently report student feedback to students (see Q4c in Table 4:14, 16 per cent, 1 or 2 rating) and the relatively low percentage of respondents saying that one of the purposes for evaluations is to feed back to students (see Q2g in Table 4:13, 44 per cent 'Yes'). Similarly, there was a low percentage of

participants saying that they seek assistance with interpreting the results from others (see Q4f in Table 4:14, 12 per cent), further supporting the outcome that evaluation seems to be a solo activity in the main. (See the discussion in section 4.3.5). The findings in relation to teacher engagement with student evaluations as expressed in the interviews differed markedly between OP and the universities. Supporting the predominance of student-focussed teaching conceptions at OP, was the fact that the majority of interviewees (17) reported using student feedback to inform and modify teaching as a routine and normative procedure. Substantially fewer interviewees at OU said they report on a systematic integration of student feedback into subsequent teaching planning (five), while four interviewees claimed they did this systematically at WU. Another 10 at WU said they engaged to some extent with student feedback in relation to their teaching.

Completing the feedback loop by reporting on student evaluations to the student body was not commonplace at any of the institutions according to the interviewees (five at OP, six at OU and seven at WU). Lecturer-student discussion at OP tended to be primarily on the course evaluations as opposed to the teacher evaluations. The timing of evaluations was raised as a key reason for not feeding back to students with even those who report discussing evaluations with students tending to do so with the next cohort (see data presented in section 4.3.5). However, it should be noted that at all the institutions some interviewees mentioned gathering feedback in ongoing or formative ways and discussing it with students.

Deliberate and systematic discussion of students' evaluation feedback does not seem to be part of the culture at any of the institutions. Discussion may occur within a course team or else tends to be of the informal, staffroom chat variety. While this was not explicitly stated by interviewees or in questionnaire comments, the timing of most evaluations at the end of the semester may also limit opportunities for post-evaluation collaborative analysis (see also section 4.3.5). The literature suggests that an important reason for the post-evaluation vacuum is that institutions tend to offer very little guidance and support around interpretation and use of evaluations (Arthur, 2009; Penny & Coe, 2004; Smith, 2008). Interviewees in the current research did not explicitly suggest that these factors were a barrier to engagement and further exploration, but this may be partly due to the fact that the questions in this regard focussed on actual behaviours rather than the reasons behind them. The main comment about interpretation surfaced from time to time in the interviews and also in the comment data from

section B of the questionnaire (see section 4.3.3.5) and comments made in response to Q21 (Theme 2 – Developmental Focus, sub-theme 2c) and Q22 (sub-theme 2d) (see section 4.3.5). These comments also pointed out the limitations of the numerical data that student evaluations/appraisals generate as opposed to qualitative comments that are received.

For the universities, this study confirmed the findings of the literature that in spite of a relatively positive attitude to evaluations, there was not a corresponding degree of engagement with, and use of, the data by academics. In this respect, the interviewees at OP expressed considerably more engagement with evaluations to inform their teaching. It is possible that interviewees did not raise the problem of interpretation and support with improvement because help in these processes has not been part of the traditional summative appraisal system. Research that involves help in interpretation and subsequent professional development (*e.g.*, Smith, 2008) needs to be extended in order to see if these additions can heighten staff engagement with evaluations and improve their usefulness for students. The culture of relative silence around student evaluations also needs further examination. The isolation of summative student evaluation systems from the rest of the teaching and learning process (Arthur, 2009) may send the unfortunate message that evaluation is an independent exercise, undertaken almost independently. This is also an important focus for future research, because until evaluation is *for* learning as well as *of* learning (Bovill, 2011), it is failing to meet the most important part of its brief, the improvement of student learning.

5.3 Summary of Chapter 5

This chapter has presented the evidence from the study in terms of a series of five assertions that, alongside the framework in Figure 5:1, highlighted the various interconnections between and amongst individual and institutional evaluation views, claims, actions and intentions. The following chapter summarises the study and presents the recommendations.

Chapter 6 : Conclusion

This purpose of this chapter is to draw the report to a close. It presents a summary of the study, and lists the recommendations and future research possibilities.

6.1 Summary of the Study

In this study, ‘student evaluation/appraisal’ referred to the standard student evaluation of teaching/course questionnaires and the associated policies, systems and processes administered centrally by institutions. The focus of the study was upon teachers and teaching, and the influence that institutional processes of student evaluation/appraisal have upon their everyday work in supporting and facilitating student learning. This focus was at the heart of the broad research question,

How do the current evaluation processes and practices influence teachers’ thinking and behaviours in relation to student learning at all stages of the teaching and learning cycle?,

and contributing sub-questions,

- 1. What views do tertiary teachers hold about student evaluations?*
- 2. What factors (causes, influences) affect these views?*
- 3. How do tertiary teachers engage with evaluation results and student feedback?*

This document is a report on this study, highlighting tertiary teachers’ perceptions of student evaluation/appraisal of teaching and courses. Tension can be felt by academic staff between and amongst: ‘ideal’ notions of student evaluation; institutional evaluation policies and processes; teacher-stated views about evaluation; actual lived experiences of evaluation; and its impact on the lives and work of academic staff. The review of the research literature in this area brought to the fore a number of issues and concerns in the area of student evaluation/appraisal. One major issue was the impact of institutional mandates on the use of standard-type student evaluation/appraisal survey instruments and systems for two different purposes, namely, for determining teaching effectiveness and quality, and for teacher professional development. While there have been claims made that the two purposes are

complementary and it is evident in some cases that efforts have been made by institutions to integrate the two purposes into the one evaluation policy and set of practices, it is how teachers *perceive* the evaluation situation within their context, and their role within that, that determines the nature and degree of their engagement with evaluation (Edström, 2008). The literature review showed that if the teacher perceptions about evaluation are not understood and acknowledged, then difficulties faced by teachers will continue and institutional expectations about evaluation will not be met.

In order to explore this felt tension, this study investigated the perceptions about evaluation/appraisal held by the tertiary teachers in three New Zealand institutions, the University of Otago (OU), The University of Waikato (WU) and the Otago Polytechnic (OP). An interpretivist research approach (Erickson, 1986) framed the study and comprised a mix of quantitative and qualitative data gathering approaches.

As part of the research frame, a literature review of research in the area of student evaluations was conducted, as well as an environmental scan of evaluation policies and practices made public by New Zealand polytechnics and universities through their websites. All teaching staff members at the three participating institutions (OU, WU and OP) were invited to complete a questionnaire to elicit their perceptions about the influence of evaluations/appraisals on teaching and learning processes from design to implementation, as well as their impact on teacher development. They were also invited to participate in an interview. In all, 2426 teaching staff were invited to respond to the questionnaire and 1065 responses were received (44 per cent). Sixty teachers (20 from each institution) were interviewed so that some deeper insights into their questionnaire responses could be gained and to provide an opportunity for those individuals to share personal experiences and viewpoints.

Analysis of the questionnaire and interview data provided the basis of a series of assertions on tertiary teachers' perceptions about student evaluation and the factors that impact upon the nature and degree of these tertiary teachers' engagement with evaluation. These assertions, presented below, were discussed in the light of the research questions, the literature review, and the original impetus for the study, which concerned experiences of working with staff on evaluation issues.

Assertion 1: *Institutional policies, processes and procedures influence teachers' perceptions about student evaluation/appraisal and their associated behaviours.*

Assertion 2: *Teachers' perceptions about the quality of student feedback, including students' ability to make judgements about teaching and courses, influence how teachers view student evaluations and engage with evaluation/appraisal data.*

Assertion 3: *The use of the same instrument for quality purposes and to inform teaching influences teachers' views of evaluation and their evaluation-related behaviours.*

Assertion 4: *Teachers' beliefs about teaching and learning and how they view themselves and their role in their institution will influence their perceptions of, and engagement with, student evaluation/appraisal.*

Assertion 5: *Teachers tend to view student evaluation/appraisal as an isolated and individual activity that informs them about their courses and provides data to demonstrate their teaching effectiveness to their institutions.*

A conceptual framework was developed to illustrate, diagrammatically, the connections between and amongst: (a) *individual perceptions* and the character of those perceptions; (b) the *practical implementation* factors surrounding the student evaluation instruments, policies and processes; and (c) the *institutional context* including expectations (*e.g.*, expressed through vision, mission and values, and more specifically through evaluation policies, processes and norms). These connections form a vital part of the process of (d) *engagement* with evaluation.

Student evaluations/appraisals involve student and institutional judgement of teachers and their professional competence. Learners are influenced by their personal and learning histories as well as their immediate experience of learning in the particular class and discipline. At this point of intersection between the person, the subject and the professional, it is inevitable that multiple factors will impinge on the way teachers conceptualise student evaluations/appraisals. As has been noted, these factors include mistrust about students' reliability, preferences for other forms of student feedback, suspicion of institutional use of student evaluations/appraisals, lack of faith in the process and instrument and suspicion of

manipulation by colleagues. There are repeated occurrences of these concerns and their prevalence varies in the different institutions.

In curriculum terms, evaluation is integral to the “artistry” of teaching and one of the key elements in artistry is an ability to critique (Eisner, 2002).

Good teaching depends upon artistry and aesthetic considerations. It is increasingly recognized that teaching in many ways is more like playing in a jazz quartet than following the score of a marching band. Knowing when to come in and take the lead, knowing when to bow out, knowing when to improvise are all aspects of teaching that follow no rule, they need to be felt. Much of good teaching is like that. (p. 382.)

Bringing this artistry to bear in courses and teaching means making decisions and applying practices that are responsive to student, curriculum and context needs. It requires being able to critique theories, actions, ideas, beliefs and practices. Eisner (2002) calls this “phronesis”, a kind of wisdom or intelligence that is integral to the art of deliberating about professional teaching and learning practices. Gathering data through a variety of means, one of which is formal student evaluation/appraisal processes, provides focus for such deliberation. Accordingly, deliberation is enhanced and enriched through contribution of perspectives beyond those of the individual teacher.

The current interest in teachers deliberating with teachers is an example of a professional practice that can refine phronesis. It can do so by creating a context where multiple interpretations and analyses are likely. Such contexts liberate one from a monocular perspective and a single interpretation. In addition, in the process teachers can strengthen their sense of community by joint deliberation. (Eisner, 2002, p. 382.)

How to create contexts within institutions that foster and support a community approach to student evaluation as described by Eisner is not a straightforward matter. While many of the issues foregrounded in this study that seem to be impeding the achievement of this vision are not new, the current research has served to highlight the experiences of New Zealand tertiary teachers in particular. The outcomes of this study thus provide a basis for recommendations about how to move towards a more shared community and collaborative view of teaching and teaching evaluation. The final phase of the study included the development of these

recommendations including suggestions for practical action in key areas. These recommendations are now presented.

6.2 Recommendations

Essentially, the recommendations below call for a concerted effort to ensure that there is a match among: conceptualisations of evaluation; how those conceptualisations are expressed in institutional policy statements and guidelines; and how they are enacted by institutions and individual teachers in their student evaluation processes and practices. Shared understandings about evaluation and student evaluation policies and practices will underpin the continuing move to developing teaching (and evaluation practices) as a collaborative active and organic endeavour, not a solitary and isolated one; an endeavour that is complementary to, not undermined by, monitoring, demonstrating and assuring quality.

The following recommendations are for institutions, the units within institutions (divisions, faculties, schools, disciplines) and for individual tertiary teachers, and are organised around a series of key areas that this study has shown to be influential in determining tertiary teachers' perceptions of student evaluation/appraisal.

Recommendation 1: *That institutions ensure that there is a clear alignment between their vision/policy statements concerning the auditing and developmental purposes of student evaluation/appraisal systems and their processes of implementation.*

This recommendation focuses on the need for institutions to decide clear purposes for their student evaluation/appraisal systems and then to ensure that there is consistency between claims made about evaluation in vision/policy documents and statements and their practical implementation. Assumptions about teaching, learning and evaluation should be inherent within policies and processes, and made explicit through their implementation.

Suggestions for achieving this include:

- Ensure that appropriate definitions of 'evaluation' underpin policy and practice. Although connected, evaluation as it applies to curriculum (which includes teaching) differs from evaluation as it applies to accountability.

- Use consistent language and terminology when referring to evaluation in all policy and resource documents, so that a uniform portrayal of the meaning/s of student evaluations is presented.
- Ensure that evaluation processes and practices enacted at all levels of the institution directly and explicitly reflect the difference between the summative and formative purposes of evaluation, as portrayed in policies and statements.
- Ensure that evaluation policies and processes align with other institutional policies such as teaching and learning plans and policies concerning staff development.
- Design student evaluation tools, instruments and processes that are fit for purpose. As Mintzberg (2004) advises,

Recall the ‘rule of the tool’ – you give a little boy a hammer and everything looks like a nail. Narrow concepts are no better than narrow techniques. Organizations [institutions and teachers] don’t need to be hit over the head with either. (p. 402.)

The use of a single instrument is not comprehensive enough to achieve both audit and development intentions. The purpose, limitations and potential of instruments should be clear, and use of data they generate should match the purpose. For example:

- divide instruments into specific areas, to meet various needs of different audiences
- include, in a formal way, a variety of student evaluation strategies, approaches and tools within the system, making clear the purposes for which each can/will be used
- review questions included in standard survey instruments regularly to ensure that they do align with updated institutional goals/plans for teaching and learning
- as a reflection of promotion and formal reviews expecting teachers to demonstrate effective teaching, promotion/review criteria should include requirements to demonstrate ways in which engagement with

evaluation/appraisal has occurred, including how response to student feedback has been undertaken.

Recommendation 2: *That institutions implement a professional development strategy that includes explicit support for the education of staff and students about the purpose of student evaluation for curriculum and teaching, and the institutional intents and purposes of its student evaluation/appraisal system.*

This recommendation highlights the need to have in place a professional development strategy that focuses on evaluation and student evaluation processes and practices. This strategy should be for staff, including teachers, administrators, heads of schools/departments and human resource personnel, and for students.

Such a strategy would be permanent and ongoing. It would be founded on a sensitivity to the many and varied meanings of evaluation that exist across any group, and a recognition that institutional policies can be interpreted in many different ways due to variations in an individual's prior and developing experience and knowledge, and the teaching, departmental/school and institutional contexts in which an individual lives and works. The strategy would position evaluation in the broader scope of curriculum, teaching and learning and student evaluation policy and practices within the wider institutional context. The education of students about evaluation would also form part of the strategy, so that they know about their roles and responsibilities in evaluation processes and how their contribution can have an impact upon themselves as learners and the quality of the learning environment.

The strategy might include a number of types of professional development activities/forms, for example:

- *embedded*, in the way 'evaluation' is portrayed through processes and practices (which will include the consistent use of language)
- *integral*, to the structure and contents of student evaluation systems, processes and practices themselves. The form the student evaluation system takes will serve to educate staff and students about the nature of evaluation. It should reflect, therefore, the assumptions about student evaluation as presented in policy

- *overt*, through workshops, training sessions and mentoring arrangements
- *in-built*, as part of normal practice, which treats evaluation as an iterative and ongoing process that includes collaborative interactions with students and other teachers (e.g., Fisher & Miller, 2008).

Recommendation 3: *That those who administer student evaluation systems recognise and acknowledge the variety of staff perceptions about student evaluation/appraisal and provide communication, support and resources that address teacher expectations and needs, without compromising institutional intents and purposes.*

This recommendation is about the need for those who devise and administer student evaluation systems to be sensitive to the perceptions of those who make use of their facilities and have to act on the outcomes generated by those systems.

It was clear from this study that the systems and processes for administering student feedback surveys have a bearing upon how staff perceive student evaluations. The study showed that there is a variety of perceptions about student evaluations held by teaching staff and a variety of reactions to participation in student evaluation systems. These perceptions were wide ranging and included clarity and confusion, comfort and discomfort, negativity, neutrality and enthusiasm. Also, the interviews provided some indication of the emotional responses that teachers experience when dealing with student evaluations, and evidence from across the data brought to light the shortfall in the ability of student evaluation systems to meet staff expectations. Whether staff perceptions about student evaluation in general or as part of institutional process/policy are accurate or erroneous is not the point. If institutions and individual teachers are to achieve their aims for student evaluation, then it is vital to acknowledge and recognise that the perceptions, whatever they are, do exist.

Suggestions for improving processes and dealing with staff perceptions that affect their view are:

Quality of the data

- provide relevant data/information in forms that are easy to access, and understandable
- provide guidance with interpretation of the data that is meaningful and useful
- provide the data within timeframes that allow the best use of the data

- be flexible with data collection timeframes to optimise response rates and provide data targeted towards improvement
- provide guidance to staff about using the data beyond their personal needs, particularly in relation to closing the feedback loop with students
- provide comparative data/benchmarks or rubrics so that staff can easily see what needs to be achieved to improve their teaching.

System design

- provide enough flexibility in systems to allow for the variety of teaching contexts/environments within the institution
- allow qualitative data collection to complement quantitative data collection methods
- provide a variety of survey methods/media to optimise data collection (*e.g.*, paper and online questionnaires)
- development targeted, easy to use resources around standard evaluation/appraisal questions
- develop a system that allows integration with other evaluation methods
- review compulsory questions for appropriateness and match with ever-evolving policy.

Education/professional development

- provide a safe and supportive environment to allow dialogue with staff, around appraisal/evaluation
- suggest specific resources or strategies to deal with shortcomings that are identified in the evaluation
- provide relevant professional development for evaluation/appraisal staff who deal with emotional and sensitive staff
- be proactive and committed to educating staff about evaluation, at the individual level, but also at departmental/school and institutional levels
- be proactive and committed to educating students about their role in evaluation

- provide comprehensive evaluation advice and information, not just in relation to student data, but also other evaluation methods that might be better in some teaching situations
- be active in institutional and national policy development in relation to evaluation of teaching and courses.

Recommendation 4: *That institutions ensure expectations about teacher and student roles and responsibilities in evaluation are unambiguous, and connections among performance, evaluation and reward are clearly understood.*

This recommendation is about the need for institutions, schools/departments/faculties and individual teachers:

- to be clear about views on evaluation
- to delineate how evaluation plays a part in monitoring and recognition of teaching performance
- to recognise and acknowledge explicitly the limitations and possibilities of an institution's student evaluation/appraisal system for facilitating processes of monitoring and reviewing teaching performance
- to ensure that the expectations of all who play a part in demonstrating and judging the quality of teaching through an institution's student evaluation/appraisal system and processes are aligned, complementary and explicit.

Examples of how this recommendation may be achieved include:

- actively promote a view that development equals quality and that evaluation underpins and facilitates the connection between development and quality
- make links between promotion decisions and the quality of evidence gathered to demonstrate teaching quality, and share with teachers, so that they gain insights into how successful teachers engage in evaluation practices and demonstrate their performance
- educate staff involved in decision making, including members of staffing advisory groups, heads of faculties/schools/departments and relevant human resource

personnel, about the limitations and possibilities of different kinds of evaluation/appraisal data for facilitating and demonstrating effective teaching

- educate staff involved in decision making, including members of staffing advisory groups, heads of faculties/schools/departments and relevant human resource personnel, in making judgements about teaching and how teaching performance can be demonstrated through the use of a variety of data, including student evaluation/appraisal data
- develop and implement protocols for communication and interaction concerning student evaluation processes and outcomes with an emphasis on transparency
- ensure that there is equity between research and teaching.

Recommendation 5: *That teachers, faculties/departments/schools and institutions embed within evaluation policies and practices the notion that a ‘well-rounded’ representation of teaching and courses is more likely to be achieved by drawing on multiple forms of evaluation data.*

In this recommendation, the focus is on the importance of multiple forms of evidence to demonstrate teaching performance. For example:

- Build into policy implementation, and, therefore, expectations, that to demonstrate effectiveness teachers should draw on a variety of data sources, such as the four types suggested by Smith (2008). In addition they should show how they make use of evaluation data as part of ongoing development. An example of how the contribution of evaluation/appraisal data to the development of a course can be communicated to relevant groups appears in Appendix 14.
- Incorporate into professional development programmes focussing on student evaluation the development of skills and understandings related to:
 - the nature of evidence in evaluation of teaching and courses
 - interpretation of data and deliberation over the meanings of findings and evidence, including the value of collaborative review and critique

- responding to data by making judgements and planning action, including communicating decisions with awareness and sensitivity to address needs of different audiences and relevant groups.
- Develop resources that include guidelines about the place and role of evaluation and practical suggestions about how to engage in evaluation for both summative and formative purposes.
- Make the connections between research and teaching clear.

Recommendation 6: *That professional development and course enhancement are firmly ensconced as the foundation and foci of student evaluation processes and practices. Institutions should devise a system that clearly defines the developmental and auditing purposes of student evaluation. The system should include processes and practices that target each purpose, but that also recognise that the purposes are complementary in nature and that a level of integration is needed to provide cohesion.*

Like most of the previous recommendations, Recommendation 6 is also about ensuring that the purposes and intents of any student evaluation system are clear to all who participate in it. Operating a single system to support both development and audit may be possible and logical in principle, because accountability and improvement are indeed bound together. However, as this study has shown, in practice, navigating one's way through what can be seen as a complicated system can be fraught. Teachers can feel that the student evaluation system limits their freedom to respond to the needs of their students and stymies their ability to demonstrate the quality of their teaching to their institution adequately. A system or systems that maps plainly and precisely how accountability and development work alongside each other, including how student evaluation plays a part, will assist in reducing tensions experienced by teaching staff.

Some suggestions include:

- Configure the student evaluation/appraisal system in the way that links audit and development within both individual and institutional student evaluation processes and practices of data collection, analysis and use. The example described by Smith (2008) provides guidance. This example includes a student evaluation framework that makes use of four different kinds of data (student experience, assessment, peer review and

personal reflection) within a five-phase model that links evaluation with staff development.

- Introduce the use of rubrics to delineate and communicate expectations of quality. Rubrics may be devised for both purposes of evaluation and used in different ways according to whether the process is a summative (audit) one or formative (developmental) one.
- Include specific elements/questions in student evaluation/appraisal instruments to meet the various needs of different audiences: teachers, institutions/departments/schools/faculties and students.
- Promote teaching as a scholarly activity that requires the application of research/inquiry approaches; with evaluation thinking being a key part of an inquiry approach.
- Create teaching networks to support collaboration, sharing and inquiry with a focus on evaluation/appraisal activity.
- Acknowledge and critique evaluation tools used, their limitations and possibilities, and capitalise upon opportunities presented by this critique for staff development, concerning data gathering, analysis, interpretation, and planning and implementing a response.
- Promote a strong culture of teaching across institutions and within institutions (divisions, faculties, schools, disciplines).
- Develop processes to reward engagement with, and responsiveness to, student evaluation feedback.

6.3 Further Investigation

There are many issues surrounding student evaluations in tertiary institutions. This study focussed particularly on teachers' perceptions of student evaluation. This particular focus meant that many other related topics were not able to be explored and indeed, during the course of the study additional avenues for investigation became evident. The following list highlights future possibilities for research.

- a) Investigate possible influences of disciplines and professions on teacher perceptions of student evaluation.
- b) Further investigate the personal and emotional aspects of involvement in evaluations.
- c) Investigate possible connections between and amongst gender, age, ethnicity and perceptions of evaluation.

In the current study, while demographic data were collected about these attributes as a way of ensuring the range of participants, analysis of possible links between the attributes and perceptions of student evaluations was beyond the scope of the study.

- d) Investigate relationship between perceptions of evaluation/appraisal and career progression.

Is there a connection between the perceptions held about evaluations/appraisal and the progression of individual through an academic career? Do those who are in positions where they have no desire to be promoted beyond their current level of employment perceive evaluations/appraisal differently from those who are keen to gain promotion?

- e) Investigate whether holding a role that includes more management than teaching, changes perceptions about evaluation held by an individual.
- f) Explore why some staff do not engage with student evaluations.

In the current study, a number of staff in each institution who do not engage with the evaluation system or who do not think involvement in student evaluations is worthwhile (83 staff across three institutions) were identified.

- g) Investigate how perceptions of student evaluations determine and influence teachers' beliefs and practices in their broader roles.

This study was based on evidence gathered from a large group of participants who provided self-reports via the questionnaire and interview. Therefore, all claims made were respondents' reports about their student evaluation views and practices. With a focus on individuals, particularly those with strong positive and negative views about student evaluation, in-depth case study investigations would extend the findings of the study to offer a level of substance and detail that this research was not able to provide.

- h) Investigate policy, practice and perspectives around ‘closing the loop’.

This study showed that teachers tended not to feedback reports and reflections on evaluation/appraisals to students, even though many of them agreed that evaluation/appraisal should benefit students and courses. Examining teachers’ perspectives of ‘closing the loop’ and their associated practices is needed to provide better understandings about this misalignment. Such investigations would inform practice and provide insight into how to generate and nurture positive beliefs, understandings and behaviours about the beneficial student learning, course enhancement and teacher professional development outcomes of ‘closing the loop’. In addition, investigations should inform how to embed notions of ‘closing the loop’ explicitly in institutional evaluation/appraisal policy, practice and process.

- i) Investigate the perceptions of evaluation/appraisal held by teachers in tertiary education organisations other than universities and polytechnics. Evaluation/appraisal in the wider tertiary education sector in New Zealand and, indeed, overseas too, is an area that is well under-researched.
- j) Finally, it would be important to investigate the views of other important groups who participate in evaluation/appraisal, namely, students and administration, management, and human resource personnel. What are the perceptions of evaluation/appraisal held by these groups and do they consider evaluation/appraisal as a worthwhile exercise? Investigation of the perspectives of other participant groups and their behaviours in relation to evaluation/appraisal would provide another important facet of insight to extend the findings of the study reported in this document.

6.4 Conclusion

Although writing about school education, the following sentiments expressed by Eisner (2004) (cited online by Smith, 2005) make sense for the tertiary sector as well:

It may be that by shifting the paradigm of education reform and teaching from one modeled after the clocklike character of the assembly line into one that is closer to the studio or innovative science laboratory might provide us with a vision that better suits the capacities and the futures of the students we teach.

As Eisner (1998) says, criticism is the process that helps perception “come into being, then later refines it and helps it to become more acute” (p. 6). But criticism can also be seen in a negative way and where teaching is concerned, negative criticism within an unclear environment does not contribute to the positive development of teaching.

Teachers do not work in isolation. They are members of departments, schools, faculties, discipline and professional groups, institutions, and the wider tertiary education sector. While all communities are connected in some way, each has its own atmosphere, culture, sets of values, aims and intentions, norms and behaviours. At each level there are multiple factors that influence and determine thinking and action. Individuals are part of the culture at each level and participate in it, enriching it, always changing it, in minor or major ways. Whatever role an individual has, it is important that an individual understands how to participate and what is expected of him or her in order to participate fully. The complexity of organisations means that learning to understand what it means to participate in a multifaceted community such as the education community, is not a straightforward venture. Learning how to participate in any community is part of being a member of that community, hence there is always learning, change and development. It is not only new members who are learning, but every member, including the well-established ‘old hands’. All members drive or cause change of some kind to happen. Communities are organic and fluid.

Where evaluation, as one activity amongst myriads of integrated activities taking place within education communities, is concerned, developing understanding about the meaning underpinning evaluation practices can be difficult and challenging for community members. Such challenges foreground the worth of the notion of evaluation by and for community members, and cause thinking, reflection, discussion and critique. Practical implementation strategies (such as survey tools to gather student responses) are devised and put in place by community members, generating activity that is typical of educational communities as a result, namely, discussion, reflection and critique. In other words, the opportunity to discuss and critique, to try and fail or succeed, are all part of the inherent and natural action that goes on within educational communities. Thus it is assumed that any aspect of core teaching and learning activity will be discussed, critiqued, experimented with, experienced, changed and developed. Evaluation is one of those core teaching and learning activities and this study, as an investigation into evaluation in the tertiary education system, was part of that ongoing

critique. One of the overarching problems uncovered by the study was that evaluating is not understood as integral to teaching and learning by all teachers.

A central vision of this project was to provide insights to assist institutions and individuals so that links could be tightened among student feedback, teaching development and student learning. It is hoped that insights gained from this project will:

- encourage a dynamic professional development environment in which teaching is critically responsive to learners
- allow a move towards a dialogical (two-way) model of teacher and student evaluation-feedback and learning-teaching, rather than a transmission based one
- enable staff and institutions, through better understanding of student evaluation, to build relationships that provide more effective teaching and a more immediate insight into student needs
- create a genuinely learner-centred educational environment in which teachers can build student understanding into the papers they teach, and students can learn about their own learning through participation in evaluative processes that help to develop a deeper understanding of the teaching *and* the learning context.

Another important facet of the study was to inform tertiary institutions about how student evaluations are being used by teaching staff. It may be timely to remind institutional decision makers of the strong developmental function of student evaluations as well as being a method of measuring quality, such as in relation to promotion. This would encourage a stronger commitment to a quality professional development culture. Institutional professional development support should always have the end goal of improving student learning through staff who have a high level of teaching skills and are able to use innovative teaching techniques to engage their students.

This study has produced empirical evidence on how evaluations are being used. It has established some next points or areas for investigation, thus providing building blocks on which current professional development systems can be critiqued and future improvements that will have more direct effects on student learning can be sought. Finding direct links between teacher professional development and student learning is not a straightforward

matter. However, creating a culture of review and evaluation focussed upon reflection on teaching and learning, with aims of enhancing learning outcomes through the appropriate use of tools such as student evaluation surveys, is a positive way forward. Consequently, recommendations from this study will help institutions and teachers to align their processes and procedures closely with student learning needs.

Appendix 1 : Questionnaire

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

Note: WU uses the term "appraisals" whereas OP & OU use the term "student evaluations"

Section A - Current Practice

The questions in Section A ask about your actual practice in relation to student evaluations/appraisals.

1 Have you ever run student evaluations/appraisals using the centralised system of evaluation/appraisal at:

1(a) Another tertiary institution?

Yes No

1(b) The University of Waikato? (WU) / Otago Polytechnic? (OP) / Your current institution? (OU)

Yes No

Even if you answered 'No' to both questions 1 and 2, we would appreciate it if you would still make general comment about your views (if any) on student evaluations/appraisals (q22) and answer the demographic information section. Thank you.

2 Please identify why you use student evaluations/appraisals:

2(a) Because it is required by my department/institution (WU & OU) / Because it is required by my school/institution (OP)

Yes No

2(b) For my own professional development

Yes No

2(c) For my promotion application

Yes No

2(d) For professional Goal Setting (PGS) purposes (WU) / For my salary review application (OP) / To meet confirmation requirements (OU)

Yes No

2(e) To help with paper refinement/documentation (WU) / To help with course refinement/development (OP & OU)

Yes No

2(f) To get feedback on my students' learning experiences

Yes No

2(g) To provide feedback to my students

Yes No

2(h) To report on quality matters to relevant internal and external bodies

Yes No

3 If there are any reasons why you use student evaluations/appraisals not covered by the above q2, please outline them here:

4 When you receive the results from your student evaluations/appraisals do you:

4(a) Spend time going over the data and responses?

Always 1 2 3 4 5 Never

4(b) Read the open question comments/responses made by the students?

Always 1 2 3 4 5 Never

4(c) Provide students with feedback on the results?

Always 1 2 3 4 5 Never

4(d) Compare the data with previous evaluations/appraisals?

Always 1 2 3 4 5 Never

4(e) Discuss the results with colleagues/teaching team?

Always 1 2 3 4 5 Never

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

4(f) Seek assistance with interpreting the results from others? (for example: colleagues, Chair of Department/School, TDU, mentor etc) (WU) / (e.g. colleagues/Head of School/Organisational Research Officer/EDC/mentor etc.) (OP) / (e.g. colleagues/Head of Dept/HEDC/mentor etc.) (OU)

Always 1 2 3 4 5 Never

4(g) Actively look for feedback about teaching and assessment?

Always 1 2 3 4 5 Never

5 If there are other things you do with your student evaluation/appraisal results, please outline them here:

6 Do you ever communicate with students about their feedback from student evaluations/appraisals?

Yes No

If you answered "No" to question 6, please go to question 9. (WU & OU)

7 Do you show your students you have taken account of their feedback from the evaluations/appraisals through:

7(a) The paper outline (WU) / The course outline (OP & OU)

Always 1 2 3 4 5 Never

7(b) Informal discussion with students

Always 1 2 3 4 5 Never

7(c) Paper refinements / improvements (WU) / Course refinements / improvements (OP & OU)

Always 1 2 3 4 5 Never

7(d) Departmental communication channels such as noticeboards, Moodle/MyWeb, email lists, etc. (WU) / School communication channels such as noticeboards, Moodle, email lists, etc. (OP) / Departmental communication channels such as noticeboards, Blackboard, email lists, etc. (OU)

Always 1 2 3 4 5 Never

8 If there are other ways you communicate with students about evaluation/appraisal results, please outline them here:

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

Section B - Perceptions

The questions in Section B ask about your perceptions of student evaluation/appraisal data and how they influence your teaching practice.

- 9 To what extent do your reasons (given in q2) for using student evaluations/appraisals influence your teaching decisions? For example: decisions about learning outcomes, teaching strategies, assessment, etc. (WU & OU) / To what extent do your reasons for using student evaluations influence your teaching decisions? For example: decisions about learning outcomes, teaching strategies, assessment, etc. (OP)
- A great deal 1 2 3 4 5 Not at all
- 10 Please explain your answer to q9: (WU & OU) / Any comments about the above question? (OP)
- 11 To what extent does your institution's use of student evaluation/appraisal data influence your teaching decisions? For example: decisions about learning outcomes, teaching strategies, assessment, etc. (WU & OU) / To what extent does Otago Polytechnic's use of student evaluation data influence your teaching decisions? For example: decisions about learning outcomes, teaching strategies, assessment, etc. (OP)
- A great deal 1 2 3 4 5 Not at all
- 12 Please explain your answer to q11: (WU & OU) / Any comments about the above question? (OP)
- 13 My paper design refinements are influenced by appraisal results: (WU) / My course design refinements are influenced by student evaluation results: (OP & OU)
- A great deal 1 2 3 4 5 Not at all
- 14 Please explain your answer to q13: (WU & OU) / Any comments about question 13? (OP)
- 15 My willingness to try new teaching approaches is constrained by the possible negative effects on my student evaluations/appraisals.
- A great deal 1 2 3 4 5 Not at all
- 16 Please explain your answer to q15: (WU & OU) / Any comments about question 15? (OP)
- 17 Do you personally consider it worthwhile to gather appraisal data about teaching and papers? (WU) / Do you personally consider it worthwhile to gather student evaluation data about teaching and courses? (OP & OU)
- Very worthwhile 1 2 3 4 5 Not at all
- 18 Please explain your answer to q17:
- 19 How effective is your institution's centralised appraisal system in gathering useful appraisal data for you? (WU) / How effective is your institution's centralised evaluation system in gathering meaningful student evaluation data for you? (OP) / How effective is your institution's centralised evaluation system in gathering useful student evaluation data for you? (OU)
- Very effective 1 2 3 4 5 Not at all
- 20 Please explain your answer to q19: (WU & OU) / Any comments about question 19? (OP)
- 21 If you were able to decide on the future of student evaluation/appraisal at your Institution, what would be your decision and why?
- 22 If you have any general comments to make about your institution's appraisal process please make them here: (WU) / Do you have any other comments to make about student evaluation of teaching/courses? (OP & OU)

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

Section C - Demographic Information

- 23 Are you: Female
Male
- 24 Which ethnic group do you associate more strongly with? (WU) / Which ethnic group do you most strongly identify with? (OP & OU)
- New Zealander of European Descent/ Pakeha*
 - Māori / New Zealander of Māori Descent
 - Pasifika
 - Asian
 - Non-New Zealand European/ Caucasian**
 - Another Ethnicity (please specify)
- Note 1: wording of two descriptors slightly different for WU (* New Zealand/Pakeha, ** European/Caucasian)*
- 25 What is your age group?
- 30 years or below
 - 31-35 years
 - 36-40 years
 - 41-45 years
 - 46-50 years
 - 51-55 years
 - 56-60 years
 - 61 years or above
- 26 How long have you been teaching in the tertiary sector?
- 0-5 years
 - 6-10 years
 - 11-15 years
 - 16-20 years
 - 21 years or more
- 27 What is the nature of your employment position at your Institution?
- Confirmation Path
 - Permanent (OU&OP) / Continuing (WU)
 - Fixed Term
 - Casual
 - Other, please specify
- 28 What position do you currently hold?
- | University of Waikato | Otago Polytechnic | University of Otago |
|-----------------------|-----------------------------------|------------------------------|
| Tutor | Lecturer | Professional Practice Fellow |
| Senior Tutor | Senior Lecturer | Teaching Fellow |
| Teaching Fellow | Principal Lecturer | Senior Teaching Fellow |
| Lecturer | Programme Manager/
Coordinator | Lecturer |
| Senior Lecturer | Head of School | Senior Lecturer |
| Associate Professor | Senior Manager | Associate Professor |
| Professor | Other, please specify | Professor |
| Other, please specify | | Other, please specify |

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

- 29 Have you APPLIED for promotion at this Institution between 2005 and 2009? (WU) / Have you APPLIED for promotion/salary review at this Institution during the previous five years? Please do not count any application made in 2010. (OP) / Have you APPLIED for promotion at this Institution between 2005 and 2009? Please do not count any application made in 2010. (OU)

Yes
No

- 30 Have you been promoted at this Institution between 2005 and 2009? (WU & OU) / Have you been promoted at this Institution during the previous five years? (OP)

Yes
No

- 31 In which Faculty/School are you based predominantly: (WU) / In which School are you based predominantly: (OP) / In which Division are you based predominantly:

University of Waikato	Otago Polytechnic	University of Otago
Arts and Social Sciences	Applied Business	Commerce
Computing and Mathematical Sciences	Architecture, Building and Engineering	Health Sciences
Education	Art	Humanities
Law	Capable NZ	Sciences
Management	Central Otago	Other, please specify
Maori and Pacific Development	Community Learning Centre	
Science and Engineering	Design	
Other (please specify):	EDC	
	Foundation Learning	
	Hospitality	
	Information Technology	
	Midwifery	
	Natural Resources	
	Nursing	
	Occupational Therapy	
	Otago Institute of Sport and Adventure	
	Social Services	
	Veterinary Nursing	
	Other, please specify	

- 32 For teaching purposes, where are you primarily based:

University of Waikato	Otago Polytechnic	University of Otago
Hamilton	Dunedin - includes Forth Street, L Block, School of Art and C Block	Invercargill
Tauranga		Dunedin
Other, please specify	Central Otago	Christchurch
	Wintec	Wellington
	Distance Lecturer based at home	Auckland
	Other New Zealand locations	Other, please specify
	Other, please specify	

Section D - Interview Availability

- 33 I am willing to participate in an interview as part of this project.

Yes
No

Appendix 2 : Research Notes Excerpt – Process of Identifying Prospective Questionnaire Respondents

- HR were asked for a list of all academic staff of University of Otago.
- They emailed us the list on 18 Feb 2010, which contained 2323 records and the fields: Title, First name, Preferred name, Surname, Email address, Department, Position title and Research Only. Chris advised that there was some duplication of employees where they have current fixed term and permanent positions. Research-only staff were identified by a “Y” in the Research Only field.
- Research-only staff were removed from the list (531x) and we removed all the duplicates (187x), which brought the number down to 1605x staff.
- As many of the email addresses were not filled in or current we had to work on the list to fill in the gaps. We cross-checked the emails with the global university email list and were able to find emails for most. 12x staff were on the global list as existing but no email address. Decided to delete these staff off the list as most appeared to be contract staff or roles that might not involve much teaching (*e.g.* Head of Discipline). Also, as so few out of such a large sample not really worth the time it would take to try and track down their emails. This left 1593 staff listed. Another 118x staff were on the list but no emails were found and they are not listed on the global system (these too were removed from the list as not possible to send invitations). 1x staff member removed as she is a member of the project team but left other HEDC staff in as felt they were removed enough from the project and would have useful perceptions to contribute from their teaching.
- The balance of staff is 1474x, which will be randomly split into three groups. This was done by adding a random number to each staff member in the spreadsheet and then sorting on that number to reorder all. The first 491 records were given the Q.ID 20100425, the next 491 records have the Q.ID 20100426 and the last 492 records have Q.ID 20100427.
- Q.ID numbers for each group are:
 - 20100425 – group one – 491x
 - 20100426 – group two – 491x

- 20100427 – group three – 492x.
- Each group was sent a slightly different survey as the order of the options for questions 2, 4 and 7 were changed to prevent bias based on how the options were ordered.
- The survey invitation was sent out on Thursday 22 Apr (0425 at 3.16pm, 0426 at 3.18pm and 0427 at 3.21pm). 3 weeks duration with two reminders.
- First reminder to be done on Monday 3 May.
- Last chance reminder to be done on Tuesday 11 May.
- Survey to be closed on Thursday 13 May.

Appendix 3 : Core Interview Questions

Interview Questions

Section 1 – Introduction questions

- a) How long have you been teaching?
- b) How much teaching do you do?
- c) Are you fully responsible for the papers/courses/programmes you teach in? How much power do you have to be able to make changes *etc.*?
- d) At what stage are you in your academic career? What are your priorities? How have they changed over time? Has the PBRF process influenced the amount of energy that you put into your teaching? Has it made any difference to how you attend to evaluations?
- e) Do you typically engage in professional development activities?

Section 2 – Views of teaching/learning/students (a-f: exploration of emotional responses; g-k: exploration of practical responses)

- a) How do you see yourself as an educator (self-concept/views of students, communication, roles *etc.*)?
- b) Generally, do you think students are able to make judgements about the quality of teaching and their learning experience?
- c) So when you receive the results of evaluations from students, how do you feel?
- d) Why do you feel that way? Or what causes you to feel that way?
- e) What do you do in response? Why?
- f) What could be done to add to the value of this experience?
- g) Have you been surprised by responses gathered through student evaluations? Why? [drawing out whether the person uses a systematic/planned approach and mix of formal & informal or not]
- h) Do you make use of the comments you get back from students via student evaluations? How?
- i) Do you discuss student feedback with colleagues (*e.g.* teaching colleagues, academic staff developers *etc.*) in order to make sense/learn from it? How does this happen? Is there encouragement to do this?
- j) Do you discuss student feedback with students? How? When, *etc.*?

Section 3 – Institutional use of evaluation (To gauge the extent to which the interviewee's formative use of the evaluation instrument for professional development is/is not affected by institutional use.)

- a) Who do you think the evaluation system is for?
- b) Follow up question: What does the institution do with the data (department, school, faculty, division, institution)?
- c) How does that impact on how you respond to evaluations and plan your teaching?
- d) Follow-up questions: Are you using it in a constructive way? Do you feel constrained?
- e) If teaching was perceived as highly as research then would the impact be different?

Section 4 – Closing questions

- a) If you had the power to change the evaluation system to maximise the value to teaching and learning, what would you do? (*e.g.* in relation to design, status, status quo, process, *etc.*).
- b) Would you like the opportunity to review the transcript of the interview?

Appendix 4 : Project Timeline

The following is an overview of project activities and their occurrence across the period of the study.

Timing	Activity
Feb – Apr 2010	<ul style="list-style-type: none"> • literature review and environmental scan and report • questionnaire tailored by each institution to suit • invitations to complete the questionnaire sent to teaching staff and prompts sent
May – Jul 2010	<ul style="list-style-type: none"> • questionnaire closed • questionnaire data collated • preliminary analysis underway
Jul – Sep 2010	<ul style="list-style-type: none"> • analysis of quantitative and qualitative questionnaire data underway • interviews arranged
Oct 2010 – Mar 2011	<ul style="list-style-type: none"> • interviews held • interview transcriptions prepared
Jan – Jun 2011	<ul style="list-style-type: none"> • comparison of quantitative data from the three institutions continued • comparison of qualitative data from the three institutions, including analysis of interview data as the interviews were completed • development of reports on qualitative and quantitative data refined and added to
May – Nov 2011	<ul style="list-style-type: none"> • comparison of combined data undertaken • reports written • review of all phases of the study • final report written (including recommendations and future action)
Oct – Nov 2011	<ul style="list-style-type: none"> • extension to project granted, work continued on final report
Nov 2011 – Jan 2012	<ul style="list-style-type: none"> • extension to project granted, work continued on final report

Appendix 5 : Questionnaire Data

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

Date Surveys Run: late April/early May 2010

Duration: three weeks

Format: on-line

Institution	Total Staff Surveyed	Total Responses	Response Rate %
Otago University (OU)	1443	670	46%
Waikato University (WU)	663	242	37%
Otago Polytechnic (OP)	320	153	48%
All Three Institutions (ALL)	2426	1065	44%

Note: WU uses the term "appraisals" whereas OP & OU use the term "student evaluations"

Section A - Current Practice

The questions in Section A ask about your actual practice in relation to student evaluations/appraisals.

1 Have you ever run student evaluations/appraisals using the centralised system of evaluation/appraisal at:

1(a) Another tertiary institution?

	Yes	No	N=	% Total Resp
OU count	173	380	553	83%
WU count	69	124	193	80%
OP count	27	80	107	70%
ALL count	269	584	853	80%
OU %	31%	69%		
WU %	36%	64%		
OP %	25%	75%		
ALL %	32%	68%		

1(b) The University of Waikato? (WU) / Otago Polytechnic? (OP) / Your current institution? (OU)

	Yes	No	N=	% Total Resp
OU count	604	56	660	99%
WU count	228	10	238	98%
OP count	124	24	148	97%
ALL count	956	90	1046	98%
OU %	92%	8%		
WU %	96%	4%		
OP %	84%	16%		
ALL %	91%	9%		

Even if you answered 'No' to both questions 1 and 2, we would appreciate it if you would still make general comment about your views (if any) on student evaluations/appraisals (q22) and answer the demographic information section. Thank you.

2 Please identify why you use student evaluations/appraisals:

2(a) Because it is required by my department/institution (WU & OU) / Because it is required by my school/institution (OP)

	Yes	No	N=	% Total Resp
OU count	462	99	561	84%
WU count	193	34	227	94%
OP count	132	4	136	89%
ALL count	787	137	924	87%
OU %	82%	18%		
WU %	85%	15%		
OP %	97%	3%		
ALL %	85%	15%		

2(b) For my own professional development

	Yes	No	N=	% Total Resp
OU count	499	71	570	85%
WU count	179	36	215	89%
OP count	115	15	130	85%
ALL count	793	122	915	86%
OU %	88%	12%		
WU %	83%	17%		
OP %	88%	12%		
ALL %	87%	13%		

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

2(c) For my promotion application

	Yes	No	N=	% Total Resp
OU count	459	106	565	84%
WU count	191	34	225	93%
OP count	70	38	108	71%
ALL count	720	178	898	84%
OU %	81%	19%		
WU %	85%	15%		
OP %	65%	35%		
ALL %	80%	20%		

2(d) For professional Goal Setting (PGS) purposes (WU) / For my salary review application (OP) / To meet confirmation requirements (OU)

	Yes	No	N=	% Total Resp
OU count	338	170	508	76%
WU count	130	73	203	84%
OP count	76	33	109	71%
ALL count	544	276	820	77%
OU %	67%	33%		
WU %	64%	36%		
OP %	70%	30%		
ALL %	66%	34%		

2(e) To help with paper refinement/documentation (WU) / To help with course refinement/development (OP & OU)

	Yes	No	N=	% Total Resp
OU count	507	62	569	85%
WU count	191	29	220	91%
OP count	128	8	136	89%
ALL count	826	99	925	87%
OU %	89%	11%		
WU %	87%	13%		
OP %	94%	6%		
ALL %	89%	11%		

2(f) To get feedback on my students' learning experiences

	Yes	No	N=	% Total Resp
OU count	543	48	591	88%
WU count	216	11	227	94%
OP count	135	6	141	92%
ALL count	894	65	959	90%
OU %	92%	8%		
WU %	95%	5%		
OP %	96%	4%		
ALL %	93%	7%		

2(g) To provide feedback to my students

	Yes	No	N=	% Total Resp
OU count	190	271	461	69%
WU count	86	116	202	83%
OP count	65	39	104	68%
ALL count	341	426	767	72%
OU %	41%	59%		
WU %	43%	57%		
OP %	63%	38%		
ALL %	44%	56%		

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

2(h) To report on quality matters to relevant internal and external bodies

	Yes	No	N=	% Total Resp
OU count	172	280	452	67%
WU count	92	103	195	81%
OP count	81	29	110	72%
ALL count	345	412	757	71%
OU %	38%	62%		
WU %	47%	53%		
OP %	74%	26%		
ALL %	46%	54%		

3 If there are any reasons why you use student evaluations/appraisals not covered by the above q2, please outline them here:

4 When you receive the results from your student evaluations/appraisals do you:

4(a) Spend time going over the data and responses?

	Always	1	2	3	4	5	Never	N=	% Total Resp
OU count	441	112	44	13	11			621	93%
WU count	138	53	32	7	5			235	97%
OP count	108	14	7	3	2			134	88%
ALL count	687	179	83	23	18			990	93%
OU %	71%	18%	7%	2%	2%				
WU %	59%	23%	14%	3%	2%				
OP %	81%	10%	5%	2%	1%				
ALL %	69%	18%	8%	2%	2%				

4(b) Read the open question comments/responses made by the students?

	Always	1	2	3	4	5	Never	N=	% Total Resp
OU count	565	33	12	6	6			622	93%
WU count	190	26	9	6	6			237	98%
OP count	124	5	3	0	3			135	88%
ALL count	879	64	24	12	15			994	93%
OU %	91%	5%	2%	1%	1%				
WU %	80%	11%	4%	3%	3%				
OP %	92%	4%	2%	0%	2%				
ALL %	88%	6%	2%	1%	2%				

4(c) Provide students with feedback on the results?

	Always	1	2	3	4	5	Never	N=	% Total Resp
OU count	33	54	87	127	302			603	90%
WU count	10	26	34	43	117			230	95%
OP count	8	26	25	29	42			130	85%
ALL count	51	106	146	199	461			963	90%
OU %	5%	9%	14%	21%	50%				
WU %	4%	11%	15%	19%	51%				
OP %	6%	20%	19%	22%	32%				
ALL %	5%	11%	15%	21%	48%				

4(d) Compare the data with previous evaluations/appraisals?

	Always	1	2	3	4	5	Never	N=	% Total Resp
OU count	330	170	68	25	21			614	92%
WU count	103	67	31	15	15			231	95%
OP count	42	41	27	11	7			128	84%
ALL count	475	278	126	51	43			973	91%
OU %	54%	28%	11%	4%	3%				
WU %	45%	29%	13%	6%	6%				
OP %	33%	32%	21%	9%	5%				
ALL %	49%	29%	13%	5%	4%				

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

4(e) Discuss the results with colleagues/teaching team?

	Always	1	2	3	4	5	Never	N=	% Total Resp
OU count		130	168	148	80	89		615	92%
WU count		34	50	70	42	35		231	95%
OP count		29	34	22	5	8		98	64%
ALL count		193	252	240	127	132		944	89%
OU %		21%	27%	24%	13%	14%			
WU %		15%	22%	30%	18%	15%			
OP %		30%	35%	22%	5%	8%			
ALL %		20%	27%	25%	13%	14%			

Note: OP has a high number of Nil responses due to a technical problem in the survey for group C (41x respondents). This affected questions 4(e) and 4(f) and all responses for group C were recorded as Nil.

4(f) Seek assistance with interpreting the results from others? (for example: colleagues, Chair of Department/School, TDU, mentor etc) (WU) / (e.g. colleagues/Head of School/Organisational Research Officer/EDC/mentor etc.) (OP) / (e.g. colleagues/Head of Dept/HEDC/mentor etc.) (OU)

	Always	1	2	3	4	5	Never	N=	% Total Resp
OU count		24	48	88	149	293		602	90%
WU count		5	13	22	53	133		226	93%
OP count		7	16	18	24	27		92	60%
ALL count		36	77	128	226	453		920	86%
OU %		4%	8%	15%	25%	49%			
WU %		2%	6%	10%	23%	59%			
OP %		8%	17%	20%	26%	29%			
ALL %		4%	8%	14%	25%	49%			

Note: OP has a high number of Nil responses due to a technical problem in the survey for group C (41x respondents). This affected questions 4(e) and 4(f) and all responses for group C were recorded as Nil.

4(g) Actively look for feedback about teaching and assessment?

	Always	1	2	3	4	5	Never	N=	% Total Resp
OU count		338	122	77	34	39		610	91%
WU count		138	41	34	10	13		236	98%
OP count		95	17	12	4	5		133	87%
ALL count		571	180	123	48	57		979	92%
OU %		55%	20%	13%	6%	6%			
WU %		58%	17%	14%	4%	6%			
OP %		71%	13%	9%	3%	4%			
ALL %		58%	18%	13%	5%	6%			

5 If there are other things you do with your student evaluation/appraisal results, please outline them here:

6 Do you ever communicate with students about their feedback from student evaluations/appraisals?

	Yes	No	N=	% Total Resp
OU count	231	399	630	94%
WU count	81	155	236	98%
OP count	76	61	137	90%
ALL count	388	615	1003	94%
OU %	37%	63%		
WU %	34%	66%		
OP %	55%	45%		
ALL %	39%	61%		

If you answered "No" to question 6, please go to question 9. (WU & OU)

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

7 Do you show your students you have taken account of their feedback from the evaluations/appraisals through:

7(a) The paper outline (WU) / The course outline (OP & OU)

	Always	1	2	3	4	5	Never	N=	% Total Resp
OU count	37	69	39	31	80			256	38%
WU count	23	34	21	10	54			142	59%
OP count	11	21	11	12	9			64	42%
ALL count	71	124	71	53	143			462	43%
OU %	14%	27%	15%	12%	31%				
WU %	16%	24%	15%	7%	38%				
OP %	17%	33%	17%	19%	14%				
ALL %	15%	27%	15%	11%	31%				

7(b) Informal discussion with students

	Always	1	2	3	4	5	Never	N=	% Total Resp
OU count	47	111	63	18	34			273	41%
WU count	22	42	33	10	34			141	58%
OP count	19	38	15	2	2			76	50%
ALL count	88	191	111	30	70			490	46%
OU %	17%	41%	23%	7%	12%				
WU %	16%	30%	23%	7%	24%				
OP %	25%	50%	20%	3%	3%				
ALL %	18%	39%	23%	6%	14%				

7(c) Paper refinements / improvements (WU) / Course refinements / improvements (OP & OU)

	Always	1	2	3	4	5	Never	N=	% Total Resp
OU count	93	136	35	7	16			287	43%
WU count	54	52	18	6	17			147	61%
OP count	34	32	2	3	1			72	47%
ALL count	181	220	55	16	34			506	48%
OU %	32%	47%	12%	2%	6%				
WU %	37%	35%	12%	4%	12%				
OP %	47%	44%	3%	4%	1%				
ALL %	36%	43%	11%	3%	7%				

7(d) Departmental communication channels such as noticeboards, Moodle/MyWeb, email lists, etc. (WU) / School communication channels such as noticeboards, Moodle, email lists, etc. (OP) / Departmental communication channels such as noticeboards, Blackboard, email lists, etc. (OU)

	Always	1	2	3	4	5	Never	N=	% Total Resp
OU count	16	35	31	39	130			251	37%
WU count	4	10	14	19	82			129	53%
OP count	4	13	11	14	19			61	40%
ALL count	24	58	56	72	231			441	41%
OU %	6%	14%	12%	16%	52%				
WU %	3%	8%	11%	15%	64%				
OP %	7%	21%	18%	23%	31%				
ALL %	5%	13%	13%	16%	52%				

8 If there are other ways you communicate with students about evaluation/appraisal results, please outline them here:

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

Section B - Perceptions

The questions in Section B ask about your perceptions of student evaluation/appraisal data and how they influence your teaching practice.

- 9 To what extent do your reasons (given in q2) for using student evaluations/appraisals influence your teaching decisions? For example: decisions about learning outcomes, teaching strategies, assessment, etc. (WU & OU) / To what extent do your reasons for using student evaluations influence your teaching decisions? For example: decisions about learning outcomes, teaching strategies, assessment, etc. (OP)

	A great deal	1	2	3	4	5	Not at all	N=	% Total Resp
OU count		97	237	142	56	37		569	85%
WU count		43	84	68	26	13		234	97%
OP count		35	42	29	5	7		118	77%
ALL count		175	363	239	87	57		921	86%
OU %		17%	42%	25%	10%	7%			
WU %		18%	36%	29%	11%	6%			
OP %		30%	36%	25%	4%	6%			
ALL %		19%	39%	26%	9%	6%			

- 10 Please explain your answer to q9: (WU & OU) / Any comments about the above question? (OP)

- 11 To what extent does your institution's use of student evaluation/appraisal data influence your teaching decisions? For example: decisions about learning outcomes, teaching strategies, assessment, etc. (WU & OU) / To what extent does Otago Polytechnic's use of student evaluation data influence your teaching decisions? For example: decisions about learning outcomes, teaching strategies, assessment, etc. (OP)

	A great deal	1	2	3	4	5	Not at all	N=	% Total Resp
OU count		59	143	168	112	87		569	85%
WU count		14	44	68	44	62		232	96%
OP count		25	46	36	17	10		134	88%
ALL count		98	233	272	173	159		935	88%
OU %		10%	25%	30%	20%	15%			
WU %		6%	19%	29%	19%	27%			
OP %		19%	34%	27%	13%	7%			
ALL %		10%	25%	29%	19%	17%			

- 12 Please explain your answer to q11: (WU & OU) / Any comments about the above question? (OP)

- 13 My paper design refinements are influenced by appraisal results: (WU) / My course design refinements are influenced by student evaluation results: (OP & OU)

	A great deal	1	2	3	4	5	Not at all	N=	% Total Resp
OU count		80	234	153	51	48		566	84%
WU count		34	86	67	27	24		238	98%
OP count		30	63	26	10	7		136	89%
ALL count		144	383	246	88	79		940	88%
OU %		14%	41%	27%	9%	8%			
WU %		14%	36%	28%	11%	10%			
OP %		22%	46%	19%	7%	5%			
ALL %		15%	41%	26%	9%	8%			

- 14 Please explain your answer to q13: (WU & OU) / Any comments about question 13? (OP)

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

- 15 My willingness to try new teaching approaches is constrained by the possible negative effects on my student evaluations/appraisals.

A great deal	1	2	3	4	5	Not at all	N=	% Total Resp
OU count	36	70	131	129	200		566	84%
WU count	11	26	44	37	117		235	97%
OP count	2	11	17	36	72		138	90%
ALL count	49	107	192	202	389		939	88%
OU %	6%	12%	23%	23%	35%			
WU %	5%	11%	19%	16%	50%			
OP %	1%	8%	12%	26%	52%			
ALL %	5%	11%	20%	22%	41%			

- 16 Please explain your answer to q15: (WU & OU) / Any comments about question 15? (OP)

- 17 Do you personally consider it worthwhile to gather appraisal data about teaching and papers? (WU) / Do you personally consider it worthwhile to gather student evaluation data about teaching and courses? (OP & OU)

Very worthwhile	1	2	3	4	5	Not at all worthwhile	N=	% Total Resp
OU count	244	154	101	44	24		567	85%
WU count	124	51	36	22	4		237	98%
OP count	91	27	14	6	1		139	91%
ALL count	459	232	151	72	29		943	89%
OU %	43%	27%	18%	8%	4%			
WU %	52%	22%	15%	9%	2%			
OP %	65%	19%	10%	4%	1%			
ALL %	49%	25%	16%	8%	3%			

- 18 Please explain your answer to q17:

- 19 How effective is your institution's centralised appraisal system in gathering useful appraisal data for you? (WU) / How effective is your institution's centralised evaluation system in gathering meaningful student evaluation data for you? (OP) / How effective is your institution's centralised evaluation system in gathering useful student evaluation data for you? (OU)

Very effective	1	2	3	4	5	Not at all effective	N=	% Total Resp
OU count	174	157	138	49	39		557	83%
WU count	88	62	50	25	10		235	97%
OP count	24	46	41	18	4		133	87%
ALL count	286	265	229	92	53		925	87%
OU %	31%	28%	25%	9%	7%			
WU %	37%	26%	21%	11%	4%			
OP %	18%	35%	31%	14%	3%			
ALL %	31%	29%	25%	10%	6%			

- 20 Please explain your answer to q19: (WU & OU) / Any comments about question 19? (OP)

- 21 If you were able to decide on the future of student evaluation/appraisal at your Institution, what would be your decision and why?

- 22 If you have any general comments to make about your institution's appraisal process please make them here: (WU) / Do you have any other comments to make about student evaluation of teaching/courses? (OP & OU)

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

Section C - Demographic Information

23 Are you:

	Female	Male	N=	% Total Resp
OU count	270	314	584	87%
WU count	124	112	236	98%
OP count	94	43	137	90%
ALL count	488	469	957	90%
OU %	46%	54%		
WU %	53%	47%		
OP %	69%	31%		
ALL %	51%	49%		

24 Which ethnic group do you associate more strongly with? (WU) / Which ethnic group do you most strongly identify with? (OP & OU)

	NZr of European Descent/ Pakeha*	Māori / NZr of Māori Descent	Pasifika	Asian	Non-NZ European/ Caucasian**	Another Ethnicity (please specify)	N=	% Total Resp
OU count	367	12	4	31	122	29	565	84%
WU count	141	16	1	11	52	18	239	99%
OP count	108	4	0	0	13	0	125	82%
ALL count	616	32	5	42	187	47	929	87%
OU %	65%	2%	1%	5%	22%	5%		
WU %	59%	7%	0%	5%	22%	8%		
OP %	86%	3%	0%	0%	10%	0%		
ALL %	66%	3%	1%	5%	20%	5%		

Note 1: wording of two descriptors slightly different for WU (* NZ/Pakeha, ** European/Caucasian)

25 What is your age group?

	30 years or below	31-35 years	36-40 years	41-45 years	46-50 years	51-55 years	56-60 years	61 years or above	N=	% Total Resp
OU count	19	43	64	79	124	113	72	74	588	88%
WU count	5	17	26	31	46	39	37	38	239	99%
OP count	3	13	20	22	23	27	12	16	136	89%
ALL count	27	73	110	132	193	179	121	128	963	90%
OU %	3%	7%	11%	13%	21%	19%	12%	13%		
WU %	2%	7%	11%	13%	19%	16%	15%	16%		
OP %	2%	10%	15%	16%	17%	20%	9%	12%		
ALL %	3%	8%	11%	14%	20%	19%	13%	13%		

26 How long have you been teaching in the tertiary sector?

	0-5 years	6-10 years	11-15 years	16-20 years	21 years or more	N=	% Total Resp
OU count	108	119	128	95	140	590	88%
WU count	32	49	56	47	55	239	99%
OP count	40	38	23	15	21	137	90%
ALL count	180	206	207	157	216	966	91%
OU %	18%	20%	22%	16%	24%		
WU %	13%	21%	23%	20%	23%		
OP %	29%	28%	17%	11%	15%		
ALL %	19%	21%	21%	16%	22%		

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

27 What is the nature of your employment position at your Institution?

	Confirmation Path	Permanent (OU&OP) / Continuing (WU)	Fixed Term	Casual	Other, please specify	N=	% Total Resp
OU count	137	396	42	5	6	586	87%
WU count	N/A	213	26	0	1	240	99%
OP count	N/A	114	12	8	0	134	88%
ALL count	137	723	80	13	7	960	90%
OU %	23%	68%	7%	1%	1%		
WU %	N/A	89%	11%	0%	0%		
OP %	N/A	85%	9%	6%	0%		
ALL %	14%	75%	8%	1%	1%		

28 What position do you currently hold?

	Professional Practice Fellow	Teaching Fellow	Senior Teaching Fellow	Lecturer	Senior Lecturer	Associate Professor	Professor	Other, please specify	N=	% Total Resp
OU count	58	26	11	90	223	84	77	13	582	87%
WU count	N/A	10	N/A	27	93	36	22	17	231	95%
OP count	N/A	N/A	N/A	51	48	N/A	N/A	0	134	88%
ALL count	58	36	11	168	364	120	99	30	947	89%
OU %	10%	4%	2%	15%	38%	14%	13%	2%		
WU %	N/A	4%	N/A	12%	40%	16%	10%	7%		
OP %	N/A	N/A	N/A	38%	36%	N/A	N/A	0%		
ALL %	6%	4%	1%	18%	38%	13%	10%	3%		

	Tutor	Senior Tutor	Principal Lecturer	Programme Manager / Coordinator	Head of School	Senior Manager	N=	% Total Resp
OU count	N/A	N/A	N/A	N/A	N/A	N/A	582	87%
WU count	3	23	N/A	N/A	N/A	N/A	231	95%
OP count	N/A	N/A	14	19	2	0	134	88%
ALL count	3	23	14	19	2	0	947	89%
OU %	N/A	N/A	N/A	N/A	N/A	N/A		
WU %	1%	10%	N/A	N/A	N/A	N/A		
OP %	N/A	N/A	10%	14%	1%	0%		
ALL %	0%	2%	1%	2%	0%	0%		

29 Have you APPLIED for promotion at this Institution between 2005 and 2009? (WU) / Have you APPLIED for promotion/salary review at this Institution during the previous five years? Please do not count any application made in 2010. (OP) / Have you APPLIED for promotion at this Institution between 2005 and 2009? Please do not count any application made in 2010. (OU)

	Yes	No	N=	% Total Resp
OU count	264	315	579	86%
WU count	154	81	235	97%
OP count	61	77	138	90%
ALL count	479	473	952	89%
OU %	46%	54%		
WU %	66%	34%		
OP %	44%	56%		
ALL %	50%	50%		

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

30 Have you been promoted at this Institution between 2005 and 2009? (WU & OU) / Have you been promoted at this Institution during the previous five years? (OP)

	Yes	No	N=	% Total Resp
OU count	268	314	582	87%
WU count	138	97	235	97%
OP count	68	71	139	91%
ALL count	474	482	956	90%
OU %	46%	54%		
WU %	59%	41%		
OP %	49%	51%		
ALL %	50%	50%		

31 In which Faculty/School are you based predominantly: (WU) / In which School are you based predominantly: (OP) / In which Division are you based predominantly:

	Commerce	Health Sciences	Humanities	Sciences	Other, please specify	N=	% Total Resp
OU count	53	243	160	120	6	582	87%
WU count	51	0	125	55	4	235	97%
OP count	13	33	42	36	8	132	86%
ALL count	117	276	327	211	18	949	89%
OU %	9%	42%	27%	21%	1%		
WU %	22%	0%	53%	23%	2%		
OP %	10%	25%	32%	27%	6%		
ALL %	12%	29%	34%	22%	2%		

Note: the different departments/schools/faculties/institutes etc of WU and OP have been grouped under the 4 OU divisions.

32 For teaching purposes, where are you primarily based:

	Invercargill	Dunedin	Central Otago	Christchurch	Wellington	Hamilton	Tauranga	Auckland	Other, please specify
OU count	2	489	N/A	44	49	N/A	N/A	0	0
WU count	N/A	N/A	N/A	N/A	N/A	230	4	N/A	4
OP count	N/A	108	9	N/A	N/A	N/A	N/A	N/A	0
ALL count	2	597	9	44	49	230	4	0	4
OU %	0%	84%	N/A	8%	8%	N/A	N/A	0%	0%
WU %	N/A	N/A	N/A	N/A	N/A	97%	2%	N/A	2%
OP %	N/A	81%	7%	N/A	N/A	N/A	N/A	N/A	0%
ALL %	0%	63%	1%	5%	5%	24%	0%	0%	0%

	Wintec	Distance Lecturer based at home	Other NZ Locations	N=	% Total Resp
OU count	N/A	N/A	N/A	584	87%
WU count	N/A	N/A	N/A	238	98%
OP count	4	10	2	133	87%
ALL count	4	10	2	955	90%
OU %	N/A	N/A	N/A		
WU %	N/A	N/A	N/A		
OP %	3%	8%	2%		
ALL %	0%	1%	0%		

Survey on Tertiary Teachers' Perceptions of Student Evaluation/Appraisal of Teaching and Courses/Papers

Section D - Interview Availability

33 I am willing to participate in an interview as part of this project.

	Yes	No	N=	% Total Resp
OU count	292	297	589	88%
WU count	121	105	226	93%
OP count	97	41	138	90%
ALL count	510	443	953	89%
OU %	50%	50%		
WU %	54%	46%		
OP %	70%	30%		
ALL %	54%	46%		

Appendix 6 : Comment Data Analysis Q10

Question Text: WU&OU: q10 - Please explain your answer to Question 9: To what extent do the reasons (given in Qu2) for using appraisals/student evaluations influence your teaching decisions (for example: decisions about learning outcomes, teaching strategies, assessment etc)? 1=A great deal, 5=Not at all.

OP: q10 - Any comments about the above question? Question 9 - To what extent do your reasons for using student evaluations influence your teaching decisions? For example: decisions about learning outcomes, teaching strategies, assessment etc. 1=A great deal, 5=Not at all.

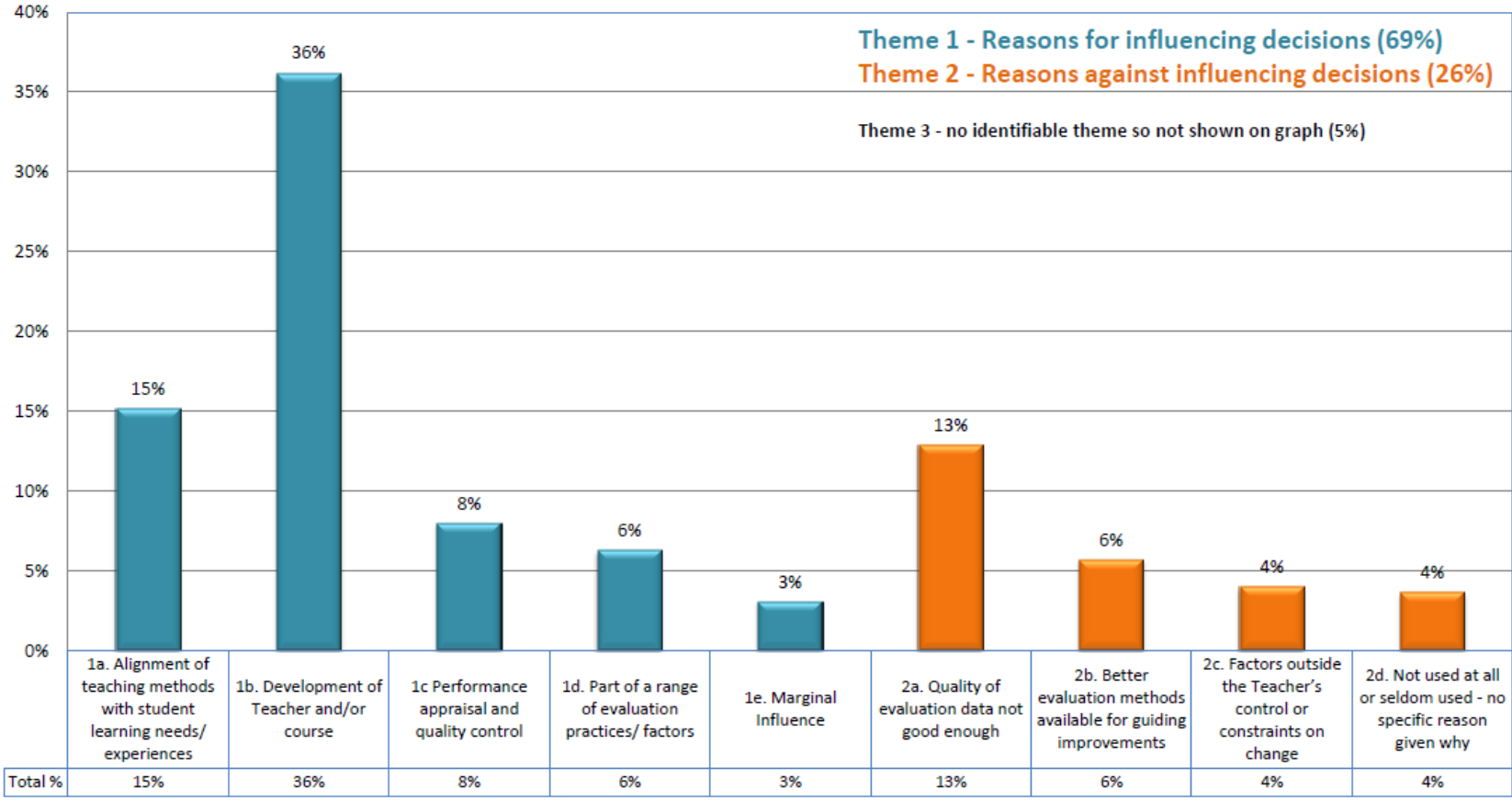
	Total Responses	Total q10 Comments	%
Otago Uni (OU):	670	472	70%
Waikato Uni (WU):	242	199	82%
Otago Poly (OP):	153	46	30%
Total:	1065	717	67%

Theme & Sub-Theme Codes for q10	Additional Detail about each sub-theme	Total Count	Total %	OU Count	OU %	WU Count	WU %	OP Count	OP %
1. Reasons for influencing decisions	Total Theme 1	570	69%	390	72%	145	61%	35	66%
1a. Alignment of teaching methods with student learning needs/ experiences	Identification of student issues, suggestions, relevance, enjoyment, effectiveness of delivery. Safe and anonymous way for students to give feedback.	126	15%	78	14%	35	15%	13	25%
1b. Development of Teacher and/or course	Teacher strengths and weaknesses, style, strategies, affirmation, professional goal setting. Course design, delivery, planning, refinement, content and assessment development. Inform future direction, test new initiatives/approaches.	301	36%	204	38%	80	34%	17	32%
1c Performance appraisal and quality control	Good evaluation results important for performance appraisal purposes and/or avoidance of teaching methods/course design that may negatively impact on evaluations. Quality control, checking or accountability aspect. Pressure to action feedback to show responsiveness.	66	8%	56	10%	10	4%	0	0%
1d. Part of a range of evaluation practices/factors	Qualitative comment data, self-reflection, assessments, informal communications with students, student reps, student responsiveness in class, informal surveys, peer review, etc.	52	6%	39	7%	12	5%	1	2%
1e. Marginal Influence	Of marginal influence or slightly useful but have definite limitations.	25	3%	13	2%	8	3%	4	8%
2. Reasons against influencing decisions	Total Theme 2	217	26%	121	22%	81	34%	15	28%
2a. Quality of evaluation data not good enough	Criticisms of instrument/process: Too limited, too broad, poor design, timing issues, prone to error, can be manipulated, not suitable for formative purposes, not appropriate for 1-to-1 or small classes, difficult to interpret as no comparisons, results received too late. Not suitable for promotion purposes as not rigorous enough. Criticisms of student feedback: Responses are contradictory, emotional, retaliatory, reaction to difficulty level, based too much on popularity/ entertainment value. Evaluations are not taken seriously by students, they are over-evaluated, it is too soon for students to understand benefits of teaching, students not qualified to comment on course content, difficulty level or recognise key skills needed.	107	13%	63	12%	37	16%	7	13%
2b. Better evaluation methods available for guiding improvements	For example: self-reflection, assessment results, informal contact/chats with students/class reps, other informal questionnaires, professional/ academic experience, teacher judgement, classroom body language.	47	6%	27	5%	17	7%	3	6%
2c. Factors outside the Teacher's control or constraints on change	Limitations on possible changes/improvements: Restrictions on time, cost, infrastructure, meeting external standards, meeting degree requirements, achieving learning objectives, nature of paper (compulsory), class size, difficult mix of students, language barriers, lack of say in paper delivery.	33	4%	18	3%	10	4%	5	9%
2d. Not used at all or seldom used (no specific reason given why)	These were generally statements which said that evaluation data was not useful or not used to inform or influence teaching practice. Only done because required. Not influenced by promotion purposes.	30	4%	13	2%	17	7%	0	0%
3. No identifiable theme	Total Theme 3	45	5%	31	6%	11	5%	3	6%
	Comments here were either not relevant (i.e. did not understand the question) or not clear enough as to meaning to be able to code.	45	5%	31	6%	11	5%	3	6%

Total All Theme Codes 832

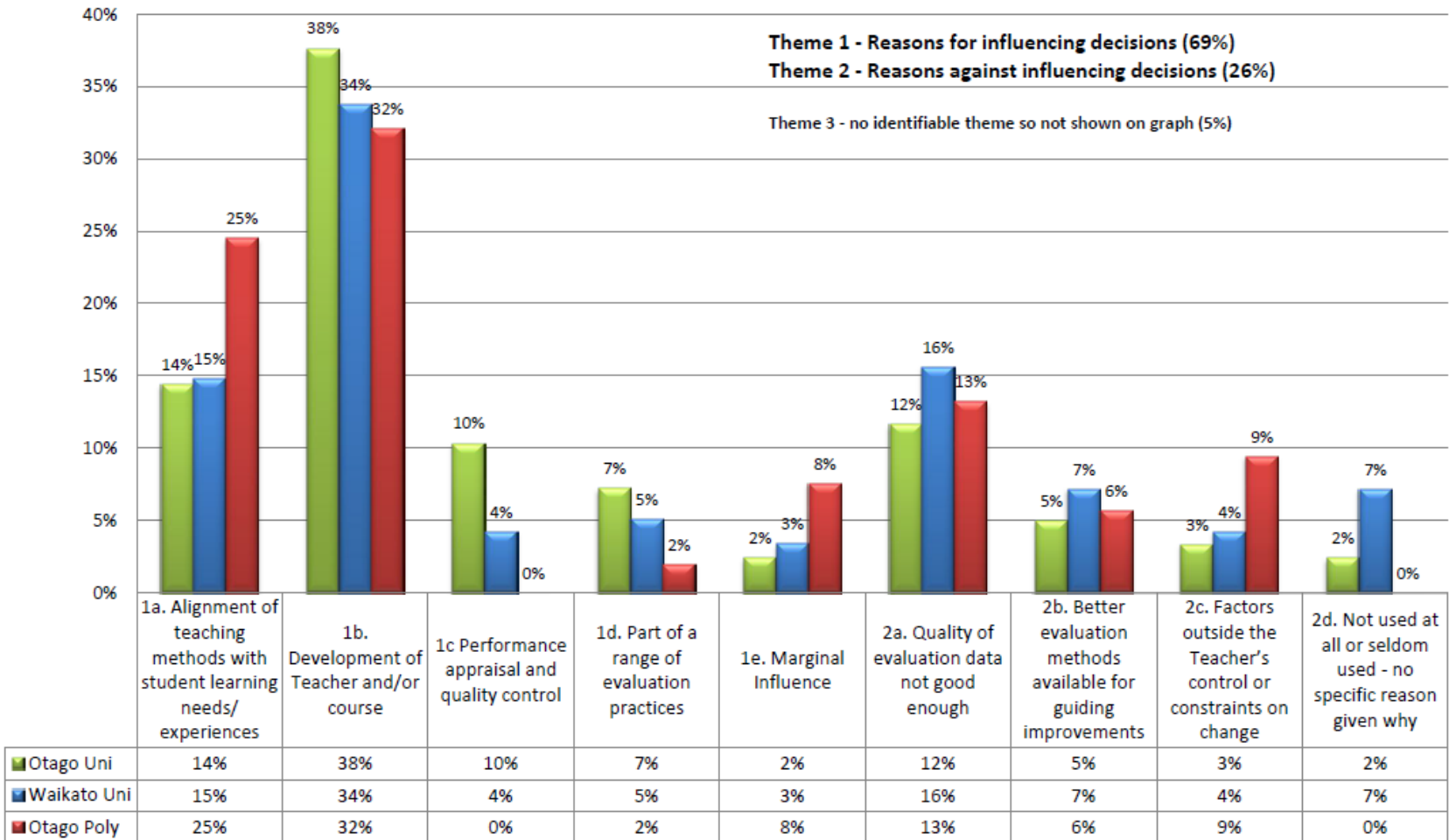
q10 Comment Data - Theme Coding - All Institutions

Question Text: q10 - Please explain your answer to Question 9: To what extent do the reasons (given in Qu2) for using appraisals/student evaluations influence your teaching decisions (for example: decisions about learning outcomes, teaching strategies, assessment etc)? 1=A great deal, 5=Not at all.



q10 Comment Data - Theme Coding - By Institution

Question Text: q10 - Please explain your answer to Question 9: To what extent do the reasons (given in Qu2) for using appraisals/student evaluations influence your teaching decisions (for example: decisions about learning outcomes, teaching strategies, assessment etc)? 1=A great deal, 5=Not at all.



Appendix 7 : Comment Data Analysis Q12

q12 Please explain your answer to Question 11 (OU & WU) / Any comments about the above question? (OP).

(OU & WU) q11: To what extent does your institution's use of student evaluation/appraisal data influence your teaching decisions? For example: decisions about learning outcomes, teaching strategies, assessment, etc 1=A great deal, 5=Not at all

(OP) q11: To what extent does Otago Polytechnic's use of student evaluation data influence your teaching decisions? For example: decisions about learning outcomes, teaching strategies, assessment, etc 1=A great deal, 5=Not at all

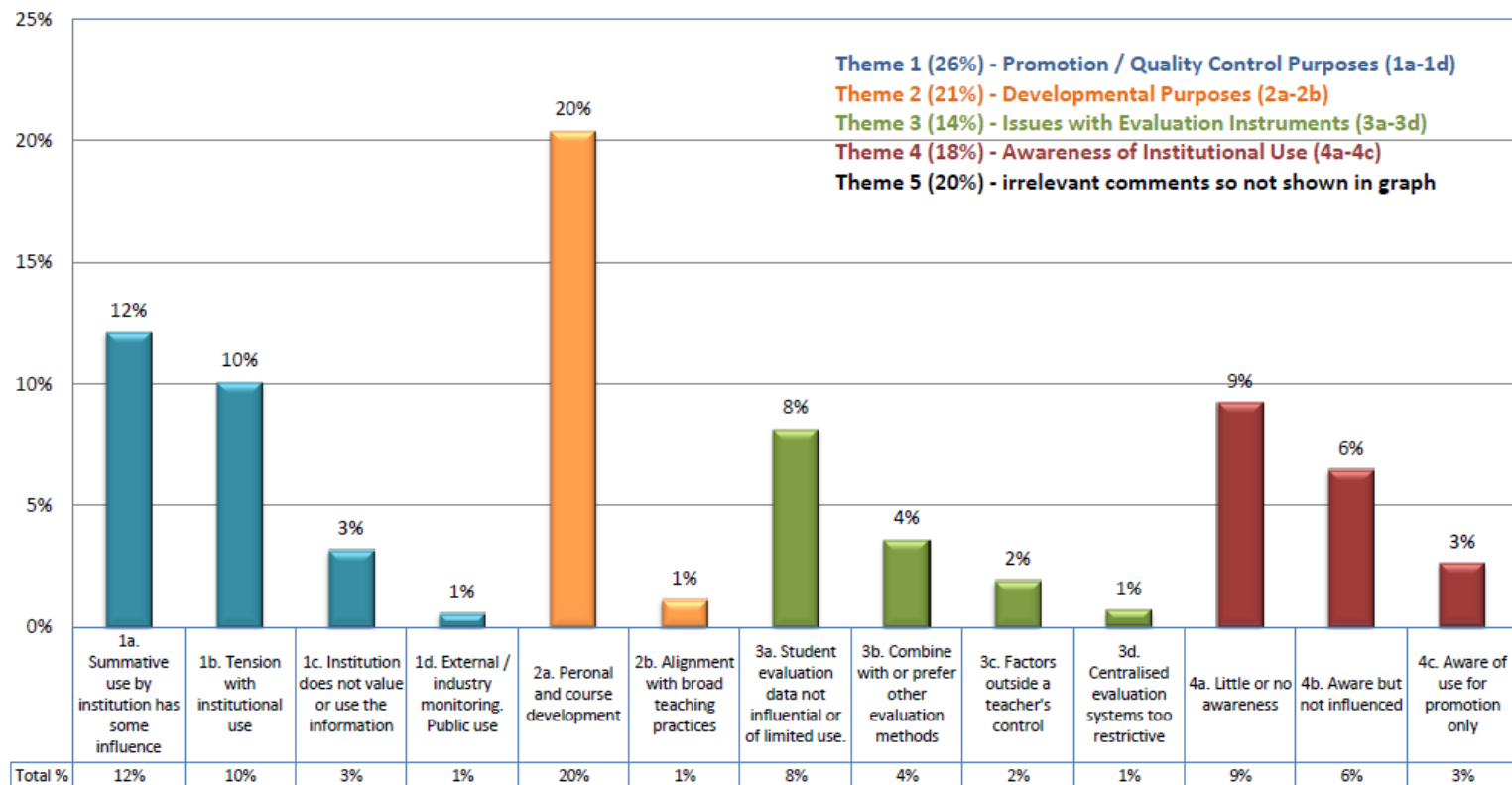
	Total Responses to Survey	Num of q12 responses	%
Otago University (OU):	670	432	64%
Waikato University (WU):	242	183	76%
Otago Polytechnic (OP):	153	39	25%
Total:	1065	654	61%
Total All Theme Codes: 726			

Theme & Sub-Theme Codes for q12	Additional Detail about each sub-theme	Total Count	Total %	OU Count	OU %	WU Count	WU %	OP Count	OP %
1. Promotion / Quality Control Purposes									
1a. Summative use by institution has some influence	Use for performance appraisal/promotions/quality assurance does have some influence teaching decisions. Need to get good ratings, awareness of promotion use, need to keep teachers on track, improve teaching. Evidence of quality. Routinely discussed with HoD/Coordinator.	88	12%	63	13%	21	10%	4	10%
1b. Tension with institutional use	Tension with/concerns about Instit use of evals/don't agree with use for performance appraisal – not appropriate use of this kind of data, Instit does not have knowledge or context of teaching environment, conflicts with formative use for own development or improving student learning, reduces innovation, trying new strategies. Can be used in a punitive way. Too much weight put on them. Can be manipulated to get good responses.	73	10%	51	11%	22	11%	0	0%
1c. Institution does not value or use the information	Instit does not do much with the information or seem to value it. No reward for excellence in results. Not as important as PBRF results/research valued more highly. Certain programmes are not valued highly by institution. Little interaction or follow-up when needed. Beaucroatic process.	23	3%	14	3%	9	4%	0	0%
1d. External / industry monitoring. Public use	Use for external monitoring/ to meet industry requirements. Should be made public so can be used for benchmarking.	4	1%	2	0%	1	0%	1	2%
Total Theme 1		188	26%	130	27%	53	26%	5	12%
2. Developmental Purposes									
2a. Peronal and course development	Use for personal/professional & course development & improving learning outcomes for students primary purpose of evals. Aids programme delivery decisions.	148	20%	103	22%	32	16%	13	31%
2b. Alignment with broad teaching practices	Enables alignment with broader teaching practices, eg colleagues, matching with graduate profiles. Aids development at departmental level.	8	1%	7	1%	1	0%	0	0%
Total Theme 2		156	21%	110	23%	33	16%	13	31%

Theme & Sub-Theme Codes for q12		Additional Detail about each sub-theme	Total Count	Total %	OU Count	OU %	WU Count	WU %	OP Count	OP %
3. Issues with Evaluation Instruments										
3a. Student evaluation data not influential or of limited use.	Student evaluation data itself not considered influential or of limited use. Criticisms of the instrument and/or the students and their ability to give feedback. Examples of criticisms include: suggestions not practical/possible, questions too limited, based on popularity/entertainment, too blunt a tool, prefer to rely on teaching experience. Only done because required to.	59	8%	36	8%	20	10%	3	7%	
3b. Combine with or prefer other evaluation methods	Combine use with or prefer other evaluation methods or to base teaching strategies on own expertise, research. Examples include: informal discussions with students, reflection, best practice, peer review, assessments, own surveys.	26	4%	15	3%	8	4%	3	7%	
3c. Factors outside a teacher's control	Factors outside the individual teacher's control eg no say over course design, content.	14	2%	10	2%	2	1%	2	5%	
3d. Centralised evaluation systems too restrictive	Restricted by centralised system of evaluation – having to do the standard one means less likely to run other, more useful evaluations.	5	1%	5	1%	0	0%	0	0%	
Total Theme 3			104	14%	66	14%	30	15%	8	19%
4. Awareness of Institutional Use										
4a. Little or no awareness	Little or no awareness of how the institution uses the data or who views evaluations/appraisals.	67	9%	35	7%	28	14%	4	10%	
4b. Aware but not influenced	Aware of institutional/performance appraisal use but no influence on teaching decisions.	47	6%	27	6%	20	10%	0	0%	
4c. Aware of use for promotion only	Only institutional use aware of is for promotion/performance appraisal.	19	3%	11	2%	8	4%	0	0%	
Total Theme 4			133	18%	73	15%	56	27%	4	10%
5. Miscellaneous										
5a. Question not answered	General comments made about evaluation/teaching strategies but question not answered.	56	8%	38	8%	12	6%	6	14%	
5b. Did not understand question	The respondents here indicated that they did not understand the question.	21	3%	12	3%	7	3%	2	5%	
5c. Referred to previous questions	Respondents referred to answer given in a different question.	55	8%	38	8%	15	7%	2	5%	
5d. Meaning not clear	The meaning of the comment not clear enough to be able to assign a theme.	13	2%	11	2%	0	0%	2	5%	
Total Theme 5			145	20%	99	21%	34	17%	12	29%

q12 Comment Data - Theme Coding - All Institutions

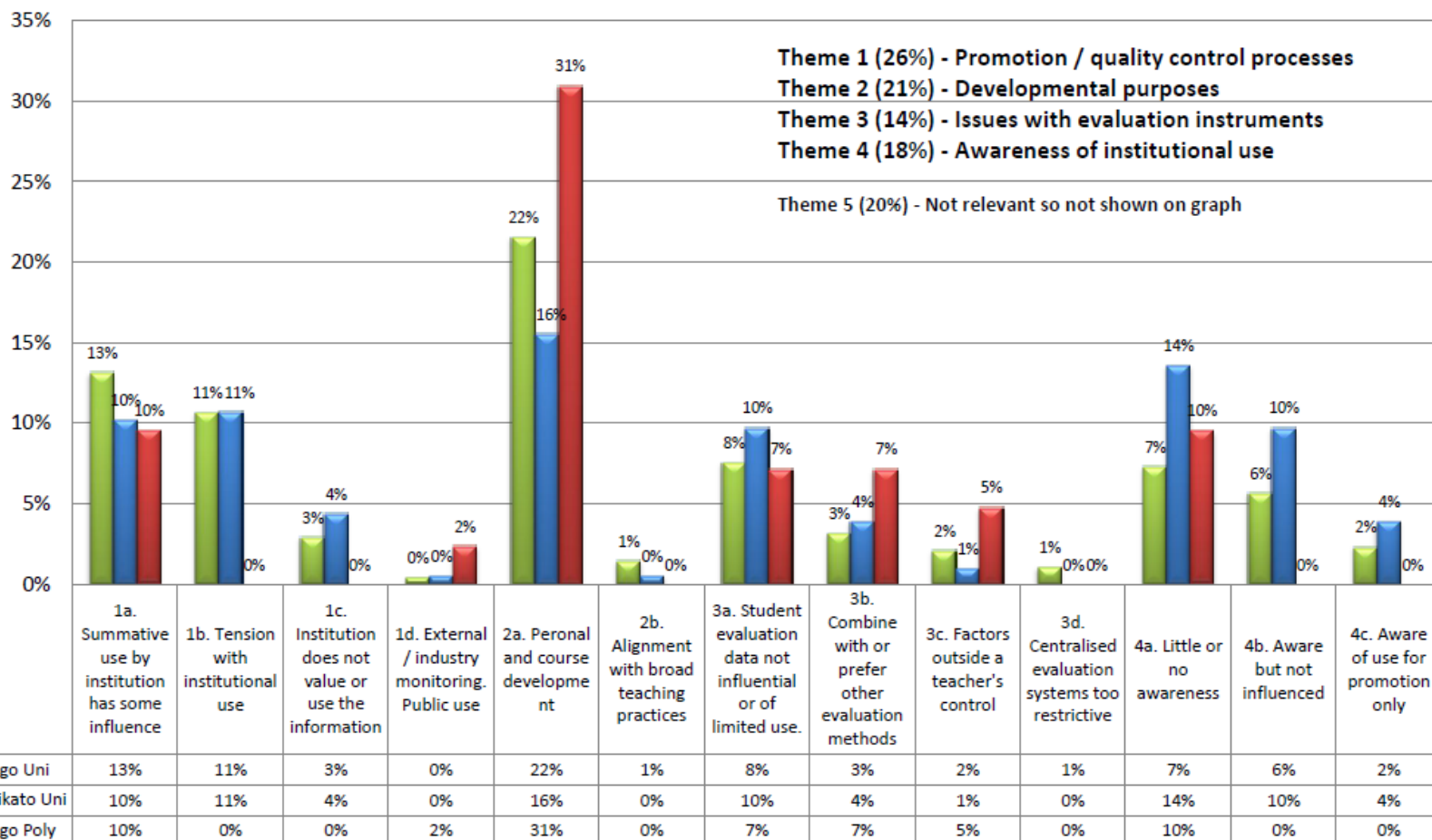
q12 Please explain your answer to Question 11 (WU&OU) / Any comments about the above question? (OP). q11: To what extent does your institution's use of student evaluation/appraisal data influence your teaching decisions (for example: decisions about learning outcomes, teaching strategies, assessment etc)? 1=A great deal, 5=Not at all



q12 Comment Data - Theme Coding - By Institution

q12 Please explain your answer to Question 11 (WU&OU) / Any comments about the above question? (OP).

q11: To what extent does your institution's use of student evaluation/appraisal data influence your teaching decisions (for example: decisions about learning outcomes, teaching strategies, assessment etc)? 1=A great deal, 5=Not at all



Appendix 8 : Comment Data Analysis Q14

q14 - Please explain your answer to q13: (OUandWU)/Any comments about question 13? (OP)

q13 - My paper design refinements are influenced by appraisal results: (WU)/My course design refinements are influenced by student evaluation results:(OUandOP)

	Total Responses to Survey	Num of q14 responses	%
Otago University (OU):	670	421	63%
Waikato University (WU):	242	182	75%
Otago Polytechnic (OP):	153	42	27%
Total:	1065	645	61%

Total All Theme Codes 706

Theme & Sub-Theme Codes for q14	Additional Detail about each sub-theme	Total Count	Total %	OU Count	OU %	WU Count	WU %	OP Count	OP %
1. Influential									
1a. Highly influential	Highly influential. Essential/Integral to the development & delivery of papers. The main purpose of collecting them. Good way to keep in touch with student thinking/expectations. Continuous improvement. Assisting student learning.	68	10%	41	9%	21	11%	6	13%
1b. Influential in specific situations	Influential in specific situations. For example: to improve teaching strategies/methods, new courses, delivery, content, readings, revising curriculum, assignments (weighting, content, timing and length), online environment. Useful for fine-tuning what is controllable. Allows to judge whether course objectives being met. Addresses where there are particular concerns by many students. Enables monitoring consistency of delivery between iterations of the same programme. Fine-tuning over a number of years.	197	28%	132	29%	53	27%	12	25%
1c. Influential alongside other evaluation practices	Influential but needs to be used in conjunction with other evaluation methods/practices. For example: self-reflection, research, peer reviews, collegial discussion, assessment grades, informal/oral feedback from students, in-class observations, informal evaluations.	47	7%	24	5%	15	8%	8	17%
1d. Comment feedback influential	Mainly the comment/qualitative feedback from the evaluations influential/used.	34	5%	21	5%	10	5%	3	6%
Total Theme 1		346	49%	218	47%	99	51%	29	60%
2. Factors reducing the amount of influence									
2a. Minimal/ no influence	Minimal influence or not influential at all. Bureaucratic process only.	20	3%	14	3%	6	3%	0	0%
2b. Student feedback of limited value	Influence reduced by limitations related to the students. For example: do not value feedback from students, students don't know enough to comment, they don't treat them seriously, they are over-evaluated, don't provide useful suggestions, provide contradictory feedback, not yet capable of valuing their learning experience, are not teaching experts, not aware of the bigger picture, students influenced by popularity/bias, negative attitude to compulsory papers. Each cohort has a very different cultural, ethnic, background etc mix so feedback from previous cohort not relevant. Students need to learn how to give constructive feedback.	51	7%	34	7%	14	7%	3	6%

q14 - Please explain your answer to q13: (OUandWU)/Any comments about question 13? (OP)

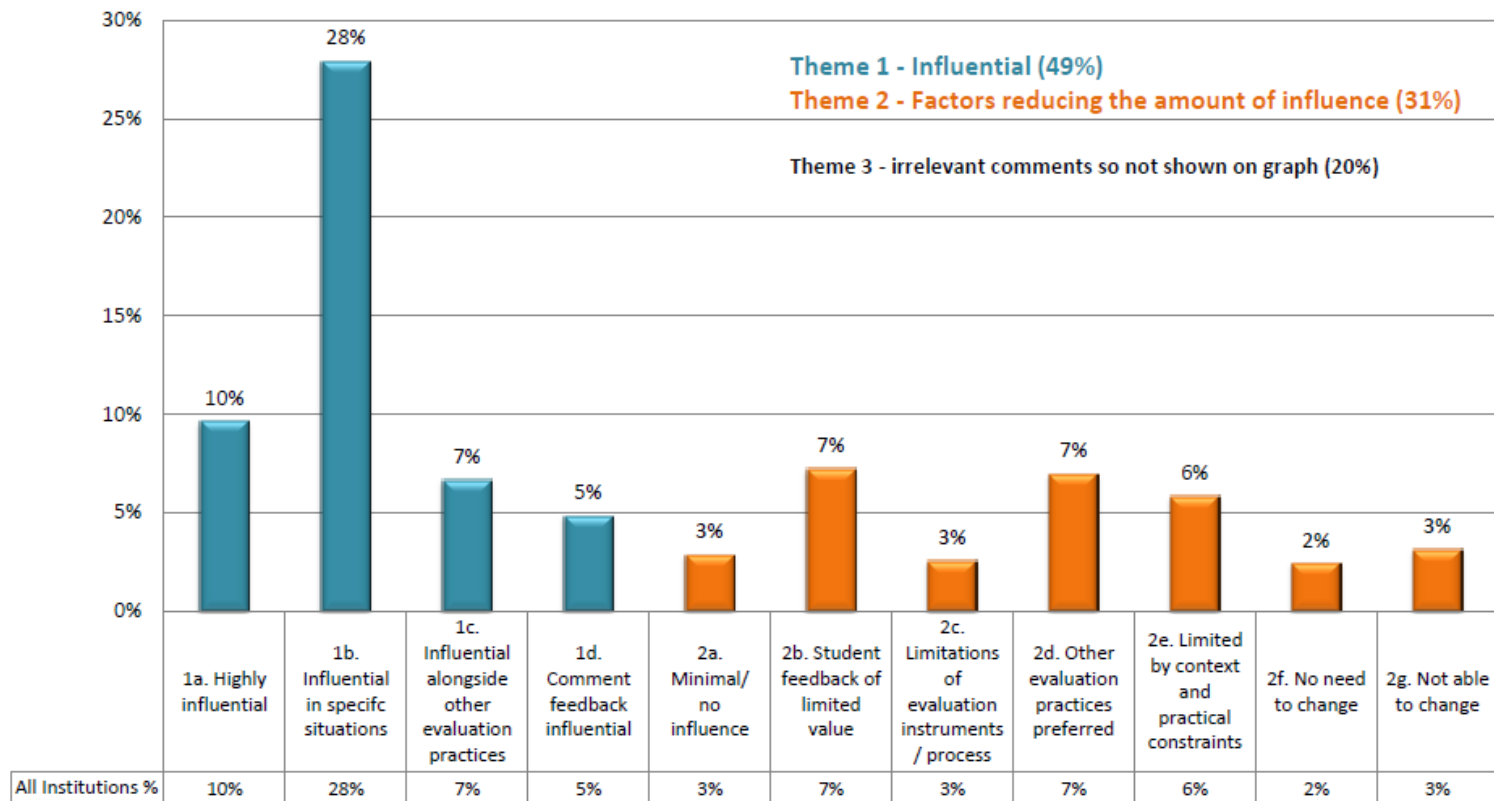
q13 - My paper design refinements are influenced by appraisal results: (WU)/My course design refinements are influenced by student evaluation results:(OUandOP)

Theme & Sub-Theme Codes for q14	Additional Detail about each sub-theme	Total Count	Total %	OU Count	OU %	WU Count	WU %	OP Count	OP %
2c. Limitations of evaluation instruments/ process	Little or no influence due to the limitations of the evaluation instruments/process. For example: delay with getting results, too blunt an instrument, ratings questions not helpful, validity questionable with small classes, questions not specific enough, mandatory questions not useful, low response rates.	18	3%	16	3%	2	1%	0	0%
2d. Other evaluation practices preferred	Other evaluation methods preferred/used. For example: self-reflection, research, peer reviews, collegial discussion, assessment grades, informal/oral feedback from students, internal review and external comment, in-class observations, better to rely on teaching experience.	49	7%	31	7%	14	7%	4	8%
2e. Limited by context and practical constraints	Amount of influence dependent on context and practical considerations/constraints. For example: clinical setting, time/workload, resourcing, curriculum limitations, external standards/requirements/accreditation. More influential with some courses than others. Need to balance what students want with what they need. Conflict between what students want and lecturer's pedagogical beliefs. Some content is required. Need to meet the needs of the discipline.	41	6%	26	6%	9	5%	6	13%
2f. No need to change	Confirmation that current approaches work so no need to change (only get good evaluations). Can only tell you if you've done well.	17	2%	12	3%	5	3%	0	0%
2g. Not able to change	Not influenced as not involved or have no say in course design or the way a course is taught.	22	3%	11	2%	8	4%	3	6%
Total Theme 2		218	31%	144	31%	58	30%	16	33%
3. Not relevant									
3a. General comments	General comments – no real reference to the question. Did not answer the question. Not able to answer the question (eg have done no evaluations to date). Did not understand the question.	19	3%	10	2%	8	4%	1	2%
3b. Referred to previous question/s	Referred back to previous question/s.	113	16%	85	18%	27	14%	1	2%
3c. Not able to code	Not clear enough as to meaning to be able to code to a theme.	10	1%	6	1%	3	2%	1	2%
Total Theme 3		142	20%	101	22%	38	19%	3	6%

q14 Comment Data - Theme Coding - All Institutions

q14 Please explain your answer to Question 13 (WU&OU) / Any comments about question 13? (OP).

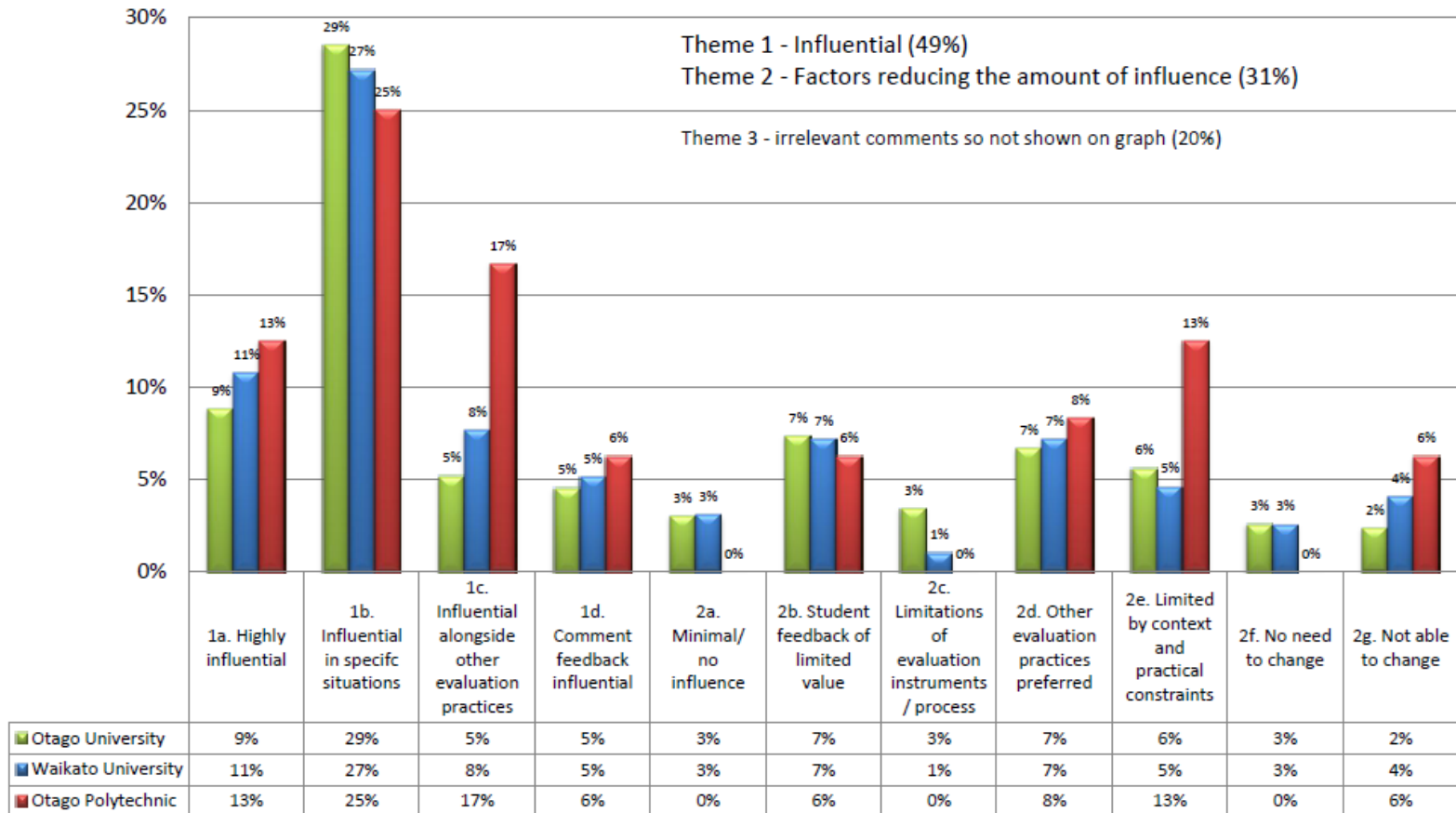
q13: My paper design refinements are influenced by appraisal results: (WU) / My course design refinements are influenced by student evaluation results: (OP & OU)



q14 Comment Data - Theme Coding - By Institution

q14 Please explain your answer to Question 13 (WU&OU) / Any comments about question 13? (OP).

q13: My paper design refinements are influenced by appraisal results: (WU) / My course design refinements are influenced by student evaluation results: (OP & OU)



Appendix 9 : Comment Data Analysis Q16

q16 - Please explain your answer to q15: (WU & OU) / Any comments about question 15? (OP)

q15 - My willingness to try new teaching approaches is constrained by the possible negative effects on my student evaluations/appraisals. (1=A great deal, 5=Not at all)

	Total Responses to Survey	Num of q16 Comments	%
Otago University (OU)	670	388	58%
Waikato University (WU)	242	174	72%
Otago Polytechnic (OP)	153	40	26%
Total All Theme Codes	706	602	57%

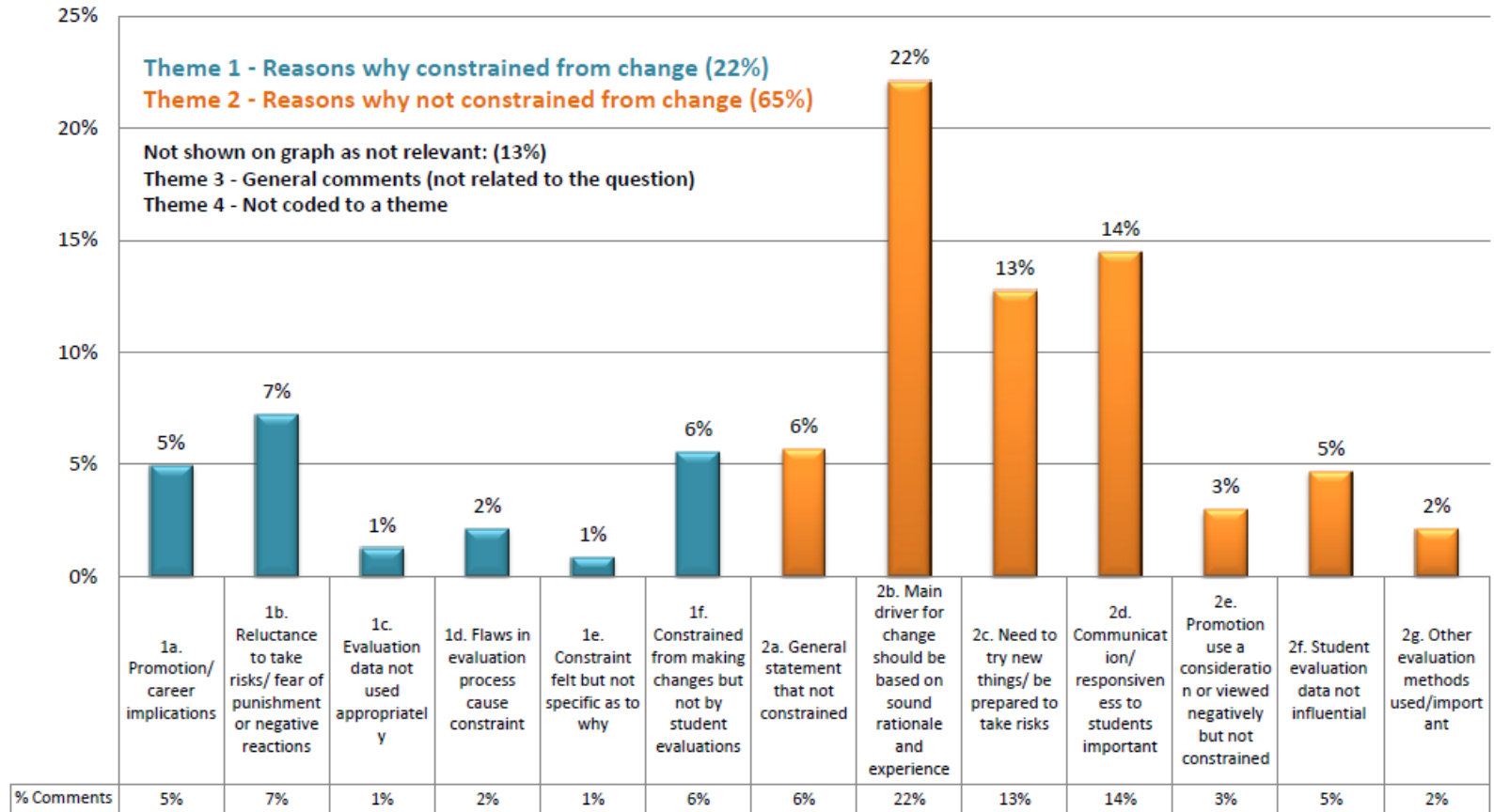
Theme Codes for q16		Total Count	Total %	OU Count	OU %	WU Count	WU %	OP Count	OP %
1. Reasons why constrained from change		155	22%	114	25%	37	18%	4	9%
1a. Promotion/ career implications	Poor evaluations damaging for promotion applications, will effect career advancement, has financial implications. More pressure to get good ratings if confirmation path/early career. Insitutional requirements, greater monitoring, competitive environment acts to constraining/increase pressure to get good ratings.	35	5%	26	6%	9	4%	0	0%
1b. Reluctance to take risks/ fear of punishment or negative reactions	Perceived need to 'play it safe' with teaching approaches. Reluctance to take risks or experiment. Don't change techniques as may effect good evaluations. Avoidance of new, innovative, demanding or difficult course material when conducting student evaluations. Feel that have to give student what they want, stay away from approaches that they don't like. Concerned about possible negative impact of an approach on evaluations. Only constrained when carrying out an evaluation, will not do a formal evaluation when trying a new method. Approach dependent on stage of career (eg more willing to take risks once confirmed). New to teaching so have less confidence/experience. Fear of poor ratings, punishment, being penalised/scrutinised. Pressure to get good ratings.	51	7%	38	8%	10	5%	3	7%
1c. Evaluation data not used appropriately	Weighted too highly, teaching/evaluation context not considered, data not interpreted correctly. Too much emphasis on evaluation data and not enough on teacher experience. Not a good way to judge teaching quality. Formative use not encouraged by process. Should not be interference from Dept/Div. Teachers should have choice what evaluation data to put forward for promotion.	9	1%	8	2%	1	0%	0	0%
1d. Flaws in evaluation process cause constraint	Criticisms of evaluation process. Focus too much on making students happy/popularity contest. Measures student happiness which is inconsistent with other UO requirements for students (eg life-long learning). Not suitable for small classes. Students expect too much. Evaluations not useful. The process is statistically invalid and unreliable.	15	2%	11	2%	4	2%	0	0%
1e. Constraint felt but not specific as to why	Mildly or somewhat constrained but not specific as to why.	6	1%	4	1%	2	1%	0	0%
1f. Constrained from making changes but not by student evaluations	Do not need to change or try new approaches as: already have good variety in course, current practices effective, prefer to stick to traditional methods, positive feedback means no need to change, don't meddle with something that works. Constrained by timetabling, resources/equipment, high workload (teaching and research), dept restrictions, conservatism of teaching culture, lack of time/opportunities to develop alternative approaches, lack of awareness/access to new teaching methods (especially in online environment). Don't have responsibility for designing courses.	39	6%	27	6%	11	5%	1	2%
2. Reasons why not constrained from change		457	65%	274	61%	146	70%	37	80%
2a. General statement that not constrained	Not constrained at all or generally. Does not occur to them to be constrained, do not see this as an issue.	40	6%	24	5%	15	7%	1	2%
2b. Main driver for change should be based on sound rationale and experience	Mostly change based on pedagogic reasons or if will enhance the learning experience. Prefer to focus on learning outcomes not evaluation outcomes. Primary aim is to fulfil course expectations. If confident teacher not likely to be constrained. Evaluations are a tool only and not decisive in determining educational aims. Prepared to live with unhappy students who benefit in retrospect from challenging methods. Will only make changes if certain they will work or not be detrimental (students are not guinea pigs). Important that changes introduced with prior investigation so response from students should not be negative. Will try new things but only if rationale well considered. If a strategy has a negative effect then not a good strategy. Recognise that new approaches need time to settle in. Long-term benefit to students, improving student learning the most important.	156	22%	93	21%	54	26%	9	20%

Theme Codes for q16		Total Count	Total %	OU Count	OU %	WU Count	WU %	OP Count	OP %
2c. Need to try new things/ be prepared to take risks	Need to be prepared to take risks, give new things a go, try and provoke students into thinking. Trying new things the only way for teaching to develop. Change is important and encourages students to think outside the square.	90	13%	49	11%	29	14%	12	26%
2d. Communication/ responsiveness to students important	Responsive to student feedback but not constrained against trying new approaches. Need students to let you know if haven't worked. See negative feedback as useful if trying new methods. Students should be encouraged to critique and comment. Actively seek feedback from the surveys about teaching approaches. Negative feedback should stimulate new approaches to change it. Use evaluations to gauge success. Good communication with students an important part of effecting change. Discuss rationale for new approach with students. Encourage students to try new activities, try and get acceptance of/engagement with change. Students support genuine attempts to improve the teaching environment.	102	14%	60	13%	32	15%	10	22%
2e. Promotion use a consideration or viewed negatively but not constrained	Use for promotion viewed negatively but not personally constrained/still willing to try new approaches. Important to try new approaches to extend teaching depth but against use in promotions. Changes should be based on academic not financial reasons. Has enough good evaluations to cover any bad ones so not constrained. Aware of promotion process and strategic about timing of evaluations but not constrained. Risk taking able to be explained in teaching portfolios so not a reason for constraint. No longer concerned about promotion. Does make more anxious but don't let it constrain.	21	3%	16	4%	5	2%	0	0%
2f. Student evaluation data not influential	Negative/low opinion of evaluations but not constrained/influenced by them. Students should not be the ones making decisions about approach/content. Delays receiving results means not constrained. Evaluations come after change has happened. See them as an administration/bureaucratic exercise only or not relevant. Not specific enough. Crude tool with results received after the fact. Students want good grades for minimum effort. Ratings data not useful. Based on popularity. Quality effected by evaluation overload. Open to manipulation.	33	5%	22	5%	7	3%	4	9%
2g. Other evaluation methods used/important	Need to use other evaluation methods as well to judge new teaching approaches. Prefer to evaluate new approaches directly and avoid using student evaluations that year. Other evaluation methods faster and more constructive. Examples of other evaluation methods: informal feedback early in semester, class discussion/reaction, student reps, one on one student comments, discussion with colleagues, review of approaches that have worked for others, peer assessment, dept student liaison group, qualitative data (eg written comments).	15	2%	10	2%	4	2%	1	2%
3. General comments		49	7%	32	7%	15	7%	2	4%
3a. Question not answered - statements about teaching approaches	General statements about teaching approaches but not specifically about evaluations or constraint. Try to mix teaching methods/refresh material to suit variety of learners. Have not changed teaching approaches.	35	5%	24	5%	9	4%	2	4%
3b. Question not answered - statements about evaluations	General statements about evaluations but does not really answer the question. Evaluations will influence the way they teach but not what they teach. Have not done or seen any evaluation results. Balance between institutional evaluation needs and teaching and learning needs.	14	2%	8	2%	6	3%	0	0%
4. Not coded to a theme		45	6%	32	7%	10	5%	3	7%
4. Not coded to a theme	Referred back to a previous question. Meaning unclear. N/A. Don't know. Too early to have an opinion. Don't understand the question.	45	6%	32	7%	10	5%	3	7%

q16 Comment Data - Theme Coding - All Institutions

q16 - Please explain your answer to q15 (WU & OU) / Any comments about question 15? (OP).

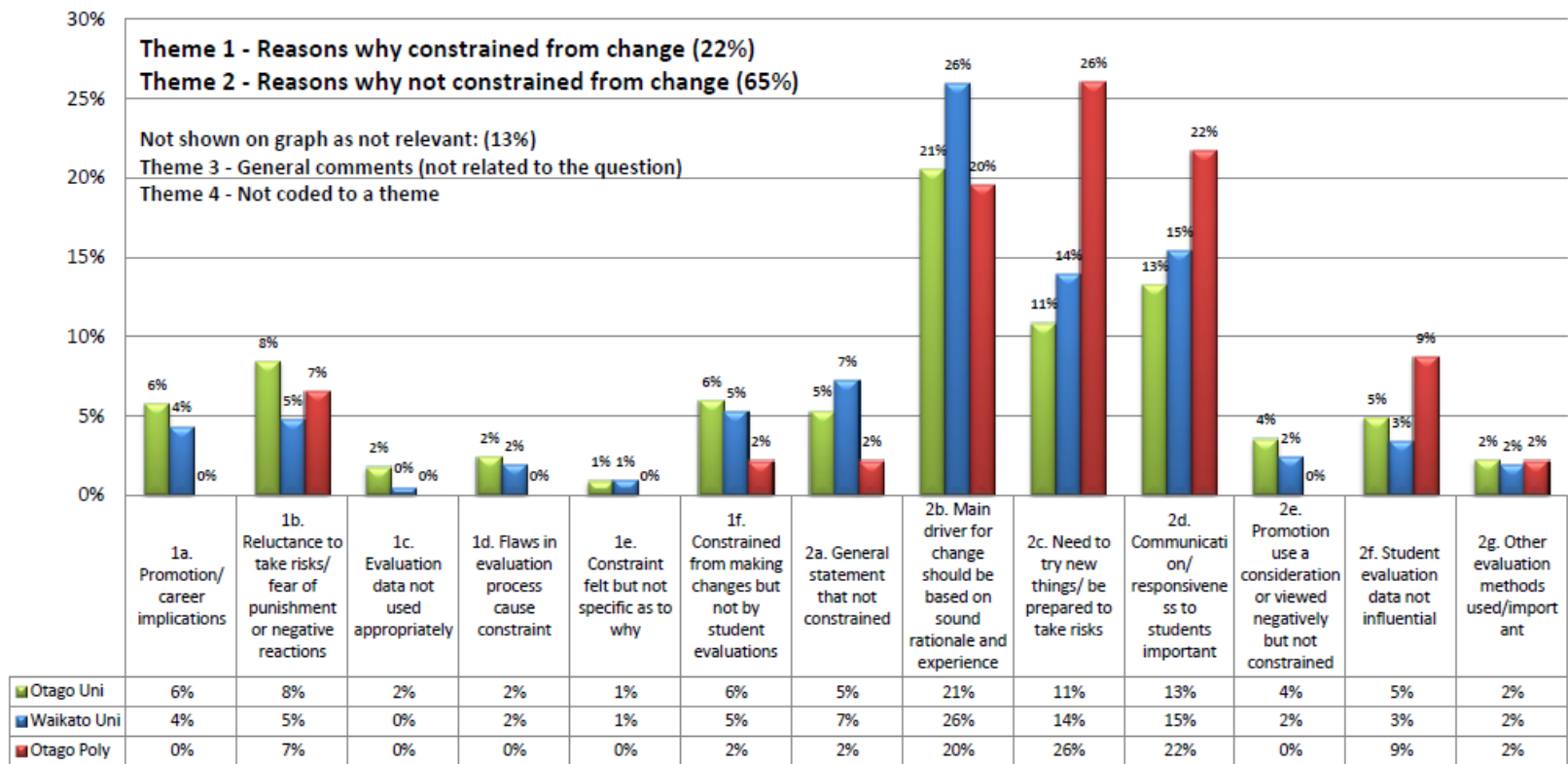
q15 My willingness to try new teaching approaches is constrained by the possible negative effects on my student evaluations/a ppraisals. 1=A great deal, 5=Not at all



q16 Comment Data - Theme Coding - By Institution

q16 - Please explain your answer to q15: (WU & OU) / Any comments about question 15? (OP).

q15 My willingness to try new teaching approaches is constrained by the possible negative effects on my student evaluations/appraisals. 1=A great deal, 5=Not at all



Appendix 10 : Comment Data Analysis Q18

q18 - Please explain your answer to Question 17:

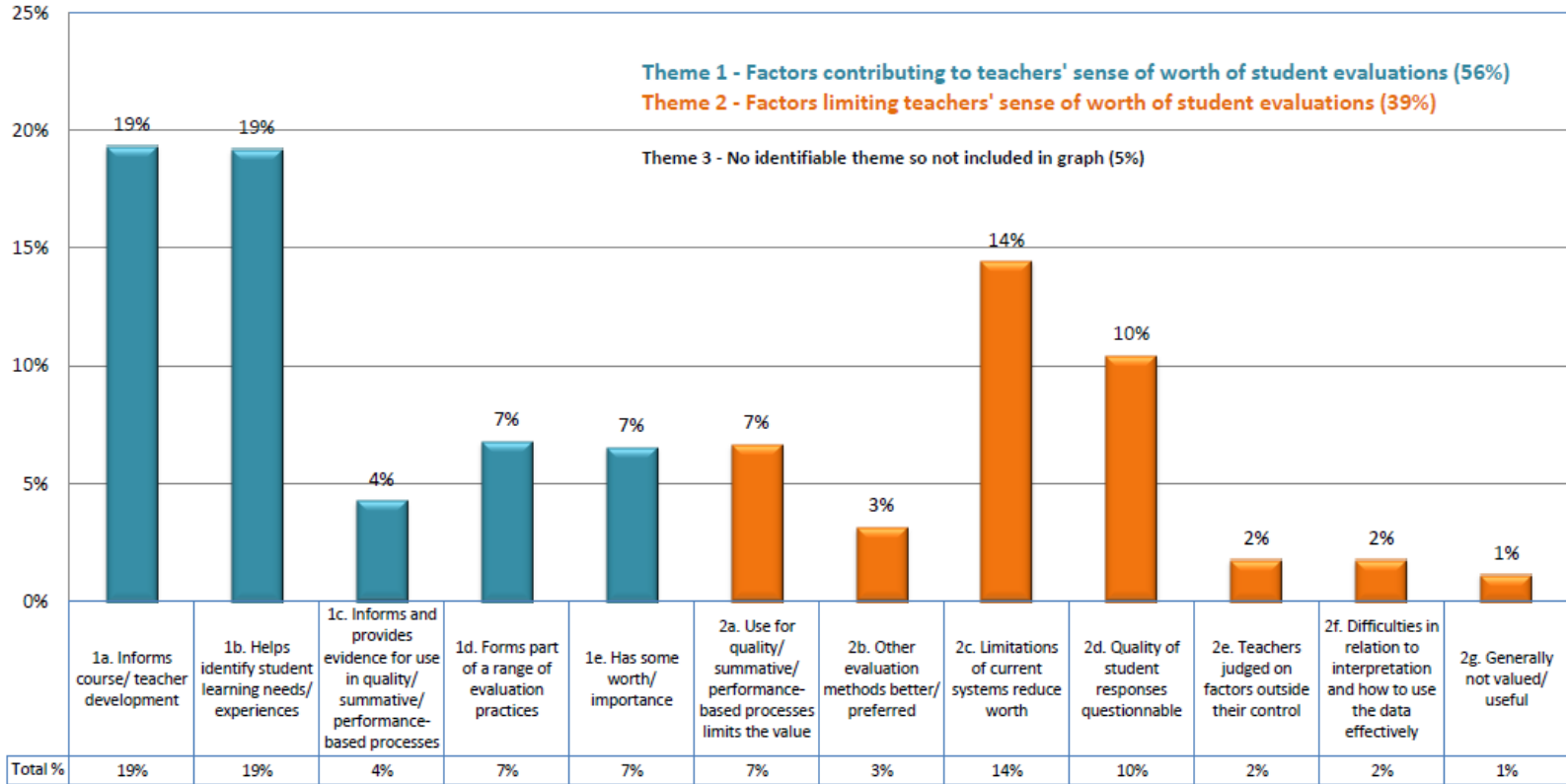
q17 - Do you personally consider it worthwhile to gather appraisal data about teaching and papers? (WU) / Do you personally consider it worthwhile to gather student evaluation data about teaching and courses? (OU & OP) 1=Very worthwhile, 5=Not at all worthwhile.

	Total Responses to Survey	Num of q18 responses	%
Otago University (OU)	670	421	63%
Waikato University (WU)	242	185	76%
Otago Polytechnic (OP)	153	68	44%
Total:	1065	674	63%
Total All Theme Codes	797		

Theme & Sub-Theme Codes for q18	Additional Detail about each sub-theme	Total Count	Total %	OU Count	OU %	WU Count	WU %	OP Count	OP %
1. Factors contributing to teachers' sense of worth of student evaluation data									
1a. Informs course/ teacher development	Teacher: skills, methods, style, strengths, weaknesses. Especially for new teachers. Course: improvement, refinement, assessment, content. Especially for new papers or when changes made. Confirmation of effectiveness/quality of teaching and courses.	154	19%	102	20%	37	17%	15	20%
1b. Helps identify student learning needs/ experiences	Gives students a voice. Part of the teaching relationship with students. Find out about student learning experiences/needs. Gauge student satisfaction. Targets the whole group not just the vocal ones.	153	19%	89	18%	50	23%	14	19%
1c. Informs and provides evidence for use in quality/ summative/ performance-based processes	Provides data for career progression/promotion, has value for institution of teaching/course quality. Has value when applying for jobs. Allows identification of teachers underperforming and needing extra support.	34	4%	16	3%	15	7%	3	4%
1d. Forms part of a range of evaluation practices	One evaluative tool of many that should be used. For example: peer review, class rep meetings, informal/passing comments from students, exam results, personal reflections, student reactions during teaching, focus groups.	54	7%	39	8%	11	5%	4	5%
1e. Has some worth/ importance	General comments that has worth but not specific as to why. Extent of importance ranges from slightly to highly. Have value to a point but have reservations.	52	7%	29	6%	14	7%	9	12%
Total Theme 1		447	56%	275	54%	127	59%	45	61%
2. Factors limiting teachers' sense of worth of student evaluation data									
2a. Use for quality/ summative/ performance-based processes limits the value	Data not sophisticated enough and of questionable validity for making judgements about teaching quality. Usefulness for formative purposes reduced by summative use and need to get positive feedback. Should not be compulsory or part of a formal process. Only used to meet requirements for promotion. Bureaucratic process.	53	7%	39	8%	13	6%	1	1%
2b. Other evaluation methods better/ preferred	Prefer to use different methods of evaluation. For example: own questionnaires, peer review, open discussion with students, exam results, student representatives.	25	3%	16	3%	6	3%	3	4%
2c. Limitations of current systems reduce worth	For example: frequency, standard nature, poor response rates, not suitable for small classes/ clinical teaching/ supervision/ prescribed courses, questionnaire format, data too crude, not focussed enough on learning outcomes, time-consuming process, lack of flexibility, impersonal, no exchange of ideas, difficult to close feedback loop due to timing/ delay with results, not relevant to area of teaching (e.g. postgraduate), questions too open to different interpretations, questions not relevant or too limited, students can't judge teacher effectiveness, qualitative (comment) data more use than quantitative data. Teacher evaluations less useful, especially for experienced, good teachers. Open to manipulation by teachers.	115	14%	81	16%	26	12%	8	11%
2d. Quality of student responses questionable	For example: don't take seriously, not well considered, contradictory responses, influenced by bias, difficulty level of course, entertainment, popularity, impractical, students lack experience and are overloaded with evaluations, confidentiality with small classes, difference between what students think important and teacher thinks important to know/learn.	83	10%	53	10%	21	10%	9	12%
2e. Teachers judged on factors outside their control	Teachers are judged by students on things outside their control. For example: staff shortfalls, timetable clashes, poor resources, no say in course design, required course content.	14	2%	7	1%	2	1%	5	7%
2f. Difficulties in relation to interpretation and how to use the data effectively	Interpretation of feedback difficult especially criticisms, unfair comments, outliers. Emotional reactions to feedback. Not sure how to use data effectively or how to close the feedback loop. Data needs to be used more effectively.	14	2%	8	2%	4	2%	2	3%
2g. Generally not valued/ useful	General comments that evaluations are not valued or useful but not specific as to why.	9	1%	4	1%	5	2%	0	0%
Total Theme 2		313	39%	208	41%	77	36%	28	38%
3. No identifiable theme									
	Comments here were either not relevant (i.e. did not understand the question), not clear enough in meaning to be able to code, or referred back to previous comments.	37	5%	25	5%	11	5%	1	1%
Total Theme 3		37	5%	25	5%	11	5%	1	1%

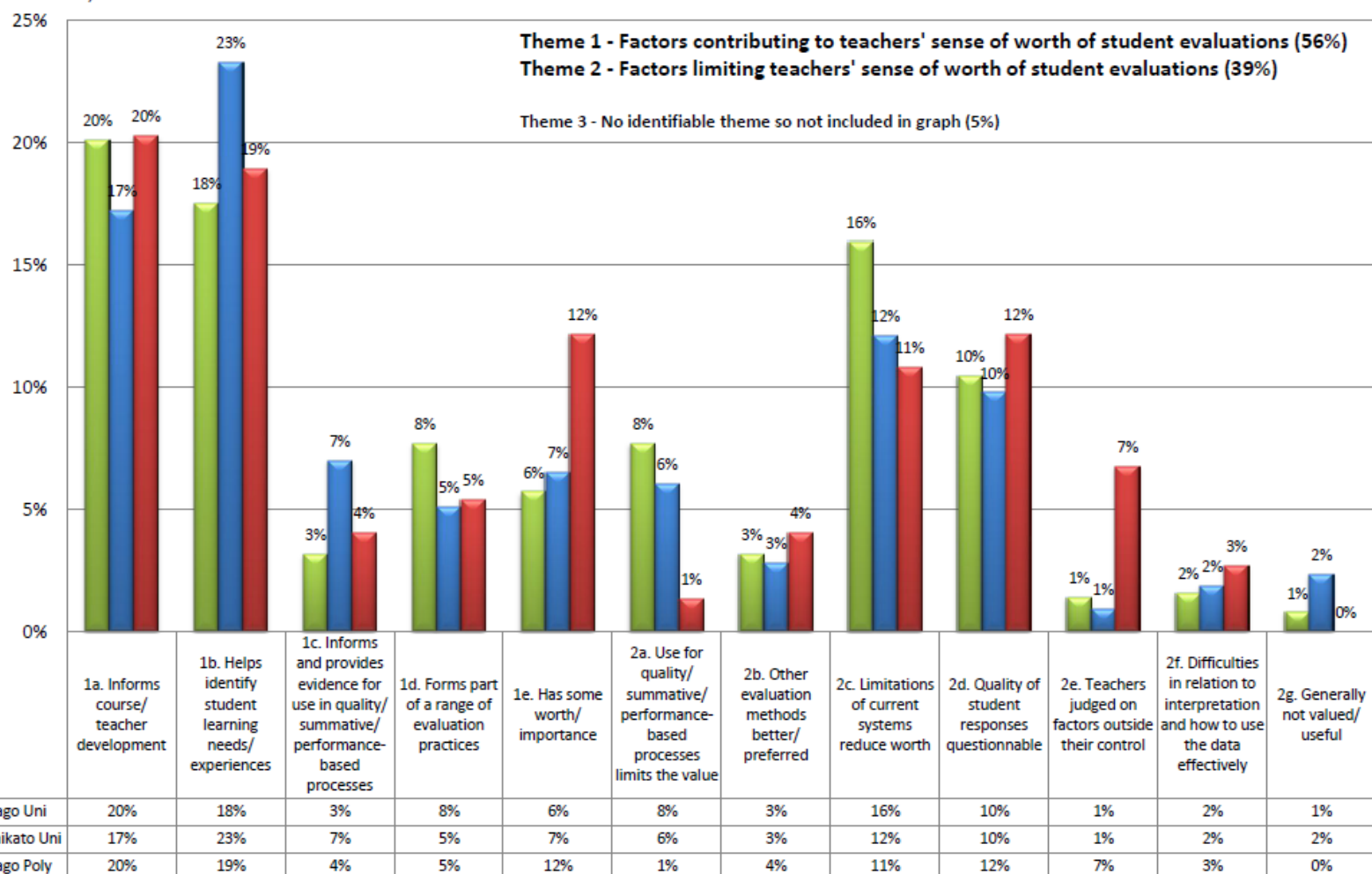
q18 Comment Data - Theme Coding - All Institutions

q18 - Please explain your answer to Question 17: Do you personally consider it worthwhile to gather appraisal/student evaluation data about teaching and papers/courses? 1=Very worthwhile, 5=Not at all worthwhile.



q18 Comment Data - Theme Coding - By Institution

q18 - Please explain your answer to Question 17: Do you personally consider it worthwhile to gather appraisal/student evaluation data about teaching and papers/courses? 1=Very worthwhile, 5=Not at all worthwhile.



Appendix 11 : Comment Data Analysis Q20

q20 Please explain your answer to q19: (WU & OU) / Any comments about question 19? (OP)

q19 How effective is your institution's centralised evaluation system in gathering useful student evaluation data for you? (OU) / How effective is your institution's centralised appraisal system in gathering useful appraisal data for you? (WU) / How effective is your institution's centralised evaluation system in gathering meaningful student evaluation data for you? (OP). 1=Very effective, 5=Not at all effective

	Total Responses to Survey	Num of q20 responses	%
Otago University (OU)	670	396	59%
Waikato University (WU)	242	179	74%
Otago Polytechnic (OP)	153	59	39%
Total:	1065	634	60%

Total All Theme Codes 767

Theme & Sub-Theme Codes for q20	Additional Detail about each sub-theme	Total Count	Total %	OU Count	OU %	WU Count	WU %	OP Count	OP %	
1. Positive aspects		Total Theme 1	347	45%	208	44%	124	57%	15	21%
1a. Gathers useful information	Comprehensive feedback. Gathers the information wanted. Useful for assisting planning/improvements, at individual level and/or departmental level. Data gathered is useful but has limitations. Generally effective/works well.	55	7%	24	5%	27	12%	4	5%	
1b. Process is efficient/ easy to do	Process is streamlined. Good to free up staff from doing their own. System is effective. Reminders to do evaluations useful. Service is good. Good support. No complaints.	210	27%	147	31%	60	27%	3	4%	
1c. Reporting/ interpretation	Easy to interpret. Reporting is clear. Analysis fast and effective. Good guidance/advice re interpretation. Easy to compare with previous evaluations.	32	4%	19	4%	10	5%	3	4%	
1d. Flexibility/ choice	Good to have choice in relation to questions, criteria, which paper to evaluate but also have access to templates.	27	4%	8	2%	17	8%	2	3%	
1e. Useful only alongside other evaluation methods	Needs to be considered along with other types of evaluation (has complementary role). E.g. informal feedback directly from students during the course, informal surveys, peer/manager review and observations, external moderation, independent focus group, class attendance and observation.	15	2%	6	1%	8	4%	1	1%	
1f. Quality control/ accountability aspects beneficial	Gives data that can be used widely. Can demonstrate to others (eg management) how the course is going. Independent. Quantifiable values useful to allow benchmarking, comparisons of trends and changes. Consistency. Centralised system necessary to maintain confidentiality.	8	1%	4	1%	2	1%	2	3%	
2. Negative Aspects		Total Theme 2	322	42%	202	43%	75	34%	45	62%
2a. Quality of student responses limited	Data obtained is too limited/blunt, not relevant/useful, important questions not answered, feedback not constructive enough. Issues relating to students assessing teaching quality: lack of knowledge, experience, biased, easily influenced by other factors such as popularity of teacher, difficulty of course, don't take seriously. Negativity of poorly performing students. Students over-evaluated so quality of responses effected. Students need guidance on how to provide constructive feedback.	70	9%	39	8%	21	10%	10	14%	
2b. System limitations	Problems with processing, e.g. delays getting back results therefore can't make immediate changes, issues with administration of survey, student attendance, poor response rate. Time-consuming/cumbersome processes. Some aspects could be easier, such as copying a previous year's evaluation. Should be able to do everything online – the requests, the surveys and have access to data and reports. Issues with statistical methods used, i.e. validity and reliability, self-selection bias. Not so useful for small classes. Lack of control, i.e. no say in what evaluations are conducted unless course coordinator, don't get to see the feedback. Evaluations should be mandated and not initiated by individual lecturer.	114	15%	84	18%	12	5%	18	25%	

q20 Please explain your answer to q19: (WU & OU) / Any comments about question 19? (OP)

q19 How effective is your institution's centralised evaluation system in gathering useful student evaluation data for you? (OU) / How effective is your institution's centralised appraisal system in gathering useful appraisal data for you? (WU) / How effective is your institution's centralised evaluation system in gathering meaningful student evaluation data for you? (OP). 1=Very effective, 5=Not at all effective

	Total Responses to Survey	Num of q20 responses	%
Otago University (OU)	670	396	59%
Waikato University (WU)	242	179	74%
Otago Polytechnic (OP)	153	59	39%
Total:	1065	634	60%

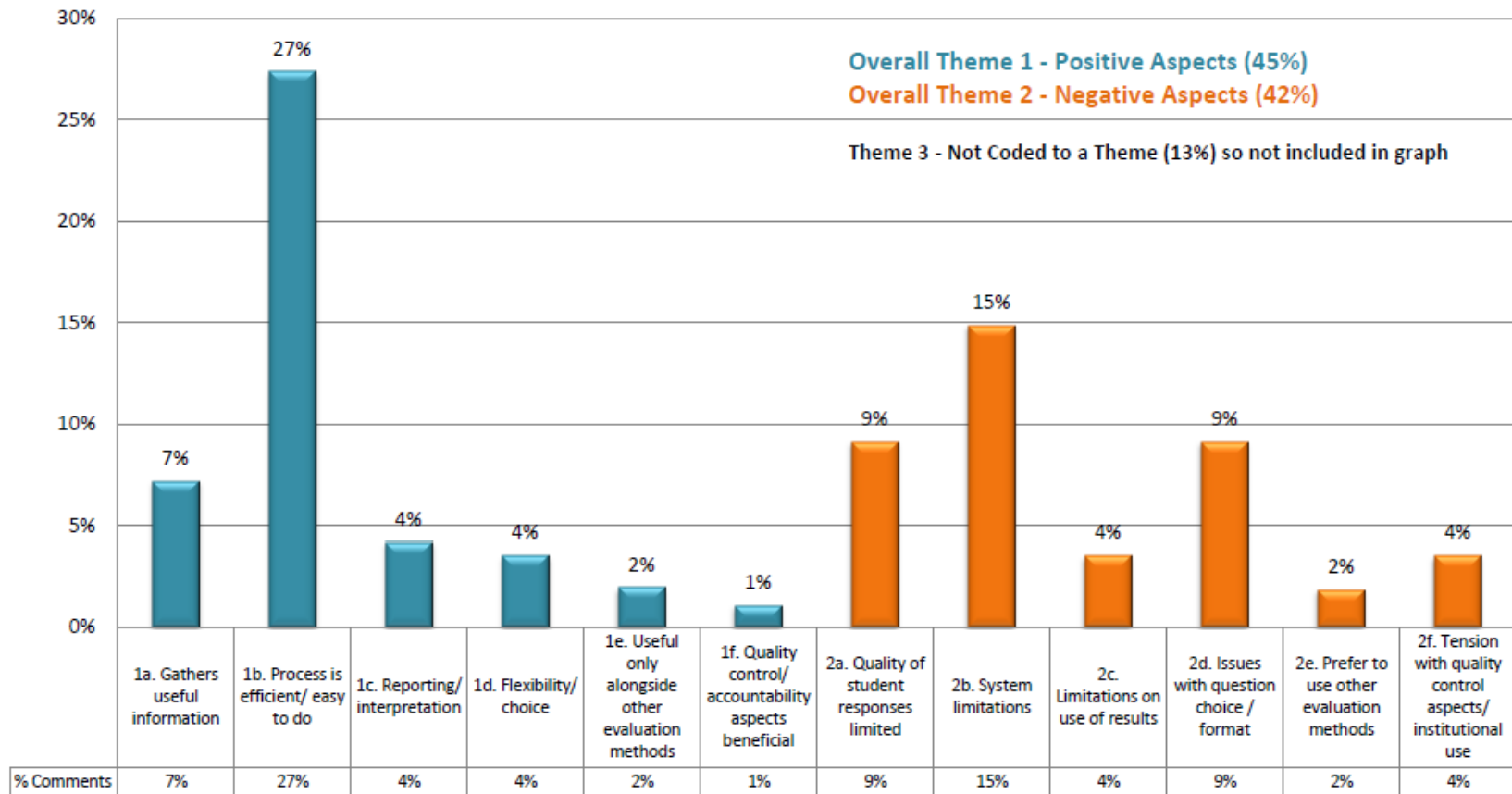
Total All Theme Codes 767

Theme & Sub-Theme Codes for q20	Additional Detail about each sub-theme	Total Count	Total %	OU Count	OU %	WU Count	WU %	OP Count	OP %
2c. Limitations on use of results	More information/guidance needed about interpretation and use of results. Not enough comparative data. Not fair to measure different teaching approaches the same way. Additional/more detailed analysis wanted, i.e. not just means, collation of comments, vetting of comments (if discriminatory). Results not used as much as they could be (negative impact on teachers). Good results not recognised/don't filter through to higher levels.	27	4%	17	4%	10	5%	0	0%
2d. Issues with question choice / format	Would like more choice/flexibility in relation to which questions to include. Questions either too generic/standardised, not individualistic enough. Questions need revision. Need a N/A option to avoid students making neutral or negative responses. Ratings questions not particularly helpful, comment question data more useful/preferred. More generic questions for specific areas, e.g. online learning. Slanted towards Sciences rather than humanities. Should be able to combine personal and course questions on same form.	70	9%	40	8%	16	7%	14	19%
2e. Prefer to use other evaluation methods	Get better information from other sources. E.g. informal feedback directly from students during the course, informal/departmental surveys, class attendance and observation.	14	2%	9	2%	5	2%	0	0%
2f. Tension with quality control aspects/ institutional use	Bureaucratic process only. Issues of academic freedom. Dual purpose discourages use for formative feedback. Open to manipulation to get a good result for promotion. Have to do them. An impersonal and automated process. Promotions committees don't know how to use them. Designed more for institutional use than individual benefit.	27	4%	13	3%	11	5%	3	4%
3. No Code									
3. Not coded to a theme	Not able to code as the meaning was unclear. Did not answer or understand the question. Do not know enough about process or haven't used much in current institution. Referred back to previous question/s.	98	13%	65	14%	20	9%	13	18%

q20 Comment Data - Theme Coding - All Institutions

q20 - Please explain your answer to q19 (OU & WU) / Any comments about question 19? (OP).

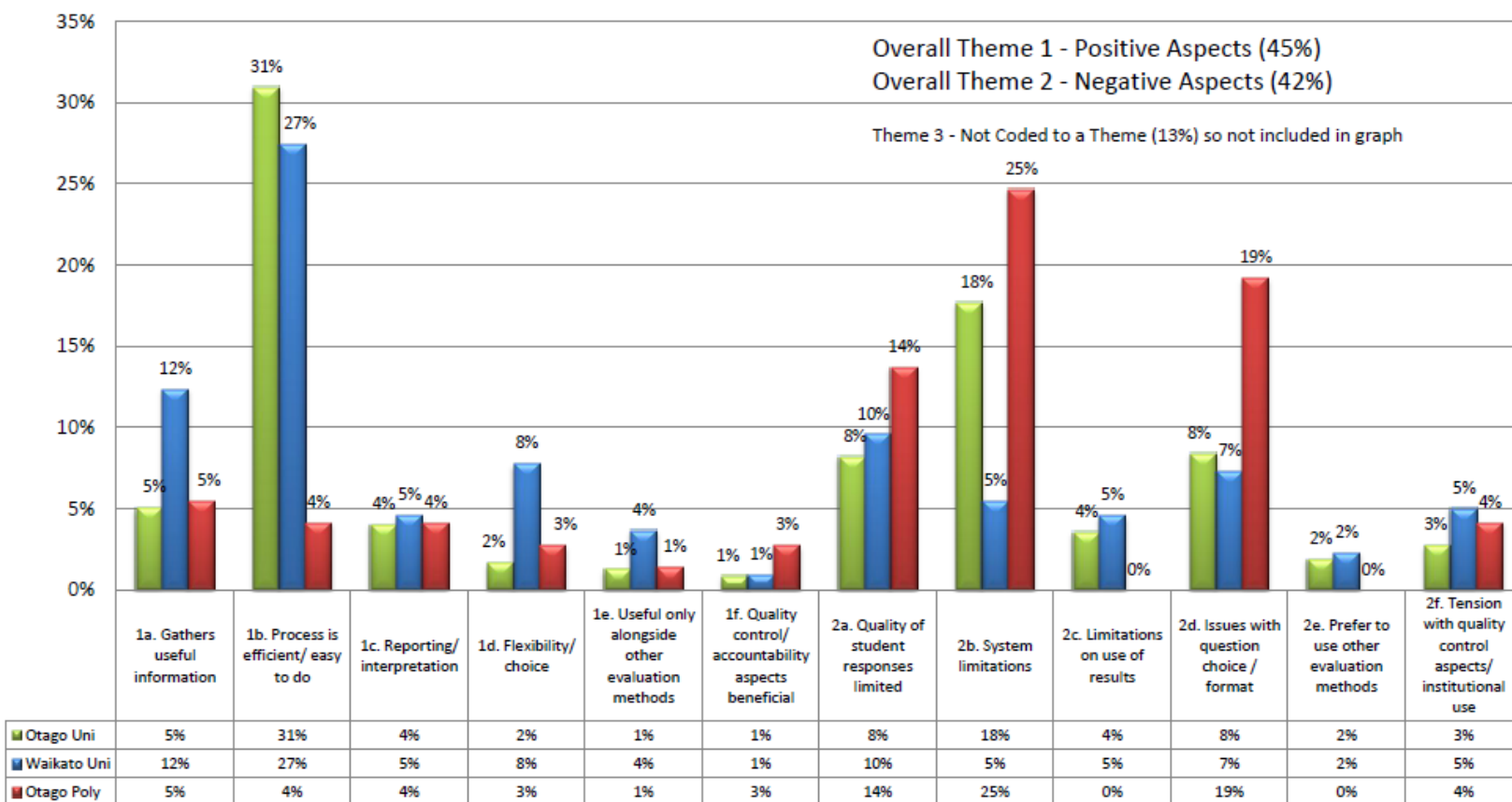
q19 How effective is your institution's centralised evaluation system in gathering useful student evaluation data for you? (OU) / How effective is your institution's centralised appraisal system in gathering useful appraisal data for you? (WU) / How effective is your institution's centralised evaluation system in gathering meaningful student evaluation data for you? (OP). 1=Very effective, 5=Not at all effective



q20 Comment Data - Theme Coding - By Institution

q20 - Please explain your answer to q19 (OU & WU) / Any comments about question 19? (OP).

q19 How effective is your institution's centralised evaluation system in gathering useful student evaluation data for you? (OU) / How effective is your institution's centralised appraisal system in gathering useful appraisal data for you? (WU) / How effective is your institution's centralised evaluation system in gathering meaningful student evaluation data for you? (OP). 1=Very effective, 5=Not at all effective



Appendix 12 : Comment Data Analysis Q21

q21. If you were able to decide on the future of student evaluation/appraisal at your Institution, what would be your decision and why?

	Total Responses to Survey	Num of q21 responses	%
Otago University (OU):	670	455	68%
Waikato University (WU):	242	204	84%
Otago Polytechnic (OP):	153	120	78%
Total:	1065	779	73%

Total All Theme Codes 1148

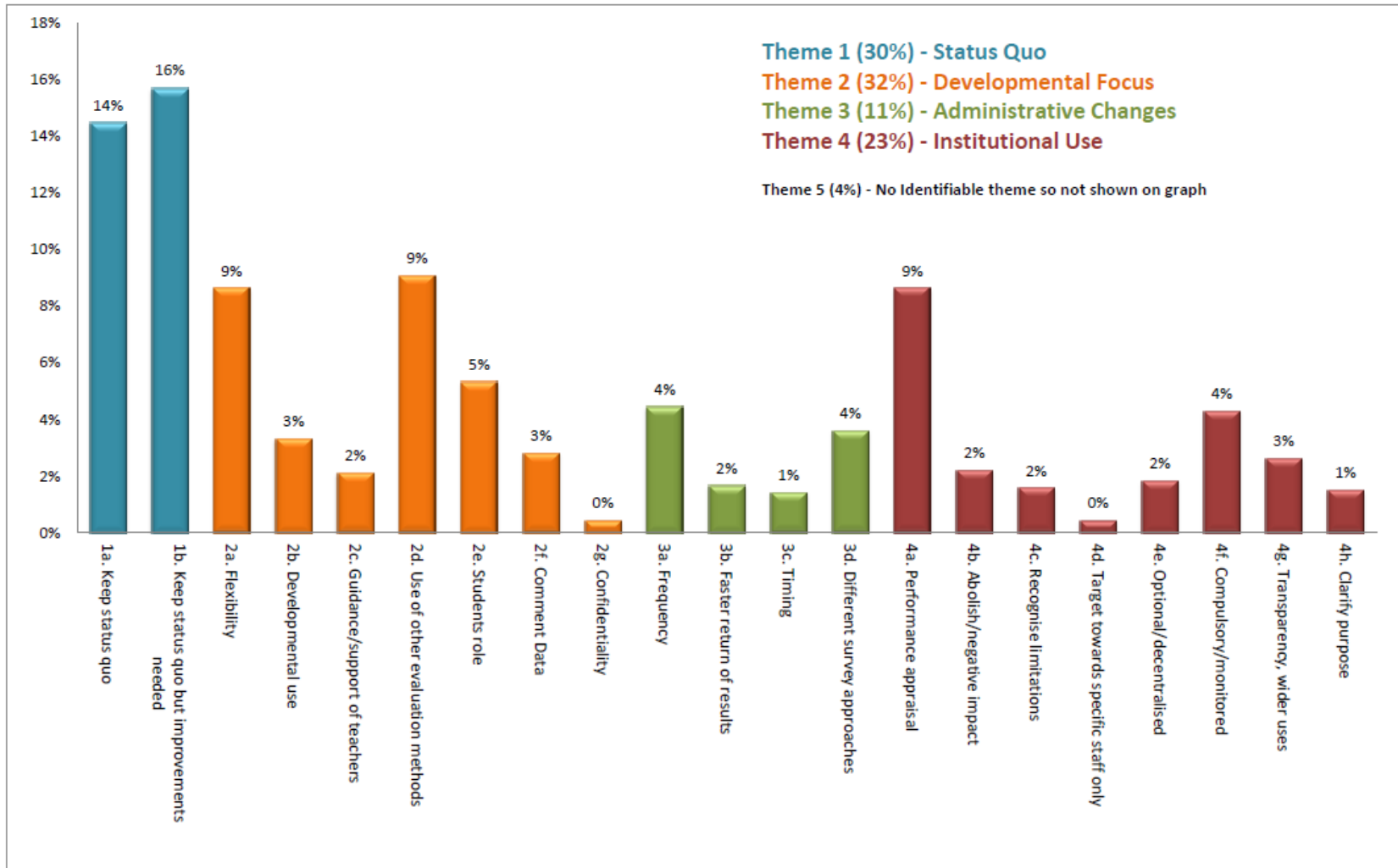
Theme & Sub-Theme Codes for q21	Additional detail about each sub-theme	Total Count	Total %	OU Count	OU %	WU Count	WU %	OP Count	OP %
1. Status quo		Total Theme 1: 346	30%	206	30%	88	30%	52	31%
1a. Keep status quo	Positive about current system, keep as is, information is important/helpful, provides evidence of quality, important for students to have a voice. Essential/must be done.	166	14%	89	13%	48	16%	29	17%
1b. Keep status quo but improvements needed	Current system is fine but some improvements suggested. Changes/review needed but no specific suggestions made.	180	16%	117	17%	40	14%	23	14%
2. Developmental Focus		Total Theme 2: 363	32%	203	30%	103	35%	57	34%
2a. Flexibility	More flexibility with the questionnaire format and questions. Questions need to be more relevant, easy to interpret, more flexible, more meaningful, broader range, specific. More guidance about good questions to ask. Provide more templates of appropriate questions. Allow teachers to design their own evaluations and have these recognised.	99	9%	57	8%	24	8%	18	11%
2b. Developmental use	More emphasis on developmental use and course evaluation for formative purposes. Simplify the process by combining course and teacher evaluation.	38	3%	27	4%	10	3%	1	1%
2c. Guidance/support of teachers	Better guidance/support for staff: how to use as a tool for improvement (ie educate staff in formative uses), how to interpret and act on the results.	24	2%	12	2%	7	2%	5	3%
2d. Use of other evaluation methods	Should be more integration with other evaluation methods, only provides partial information. Other evaluative methods are better for development and to assess quality. E.g. peer & expert review, class reps, informal dialogue, informal surveys/quizzes, focus groups, student performance, reflective journalling, external assessment, etc. Need better evaluation for graduate level.	104	9%	62	9%	32	11%	10	6%
2e. Students role	Educate students about value and consequences, better buy-in from students needed. Need to help the students provide more useful information. Close the feedback loop so students treat more seriously. Issues with student data including bias, personality/popularity effects, reaction to being challenged or difficulty level, low response rates, not evaluating all the students. Should not be anonymous, students should be accountable for their responses.	61	5%	28	4%	15	5%	18	11%
2f. Comment Data	More emphasis on the qualitative data from open-ended comment questions as more value, need to encourage more of this type of feedback.	32	3%	16	2%	12	4%	4	2%
2g. Confidentiality	Should be confidential and private to the teacher.	5	0%	1	0%	3	1%	1	1%

q21. If you were able to decide on the future of student evaluation/appraisal at your Institution, what would be your decision and why?

Theme & Sub-Theme Codes for q21		Additional detail about each sub-theme	Total Count	Total %	OU Count	OU %	WU Count	WU %	OP Count	OP %	
3. Administrative changes			Total Theme 3:	127	11%	85	12%	20	7%	22	13%
3a. Frequency		Change frequency of evaluations. Some staff want less, some want more. Many concerns with over-evaluation and the effect this has on quality of the responses.	51	4%	37	5%	4	1%	10	6%	
3b. Faster return of results		Need results returned faster.	19	2%	13	2%	2	1%	4	2%	
3c. Timing		Timing of evaluation important. Earlier is better so changes can be made and data is more useful, difficult to avoid assessment activities which influence quality of results. Should evaluate the following year to allow comparison with other papers in course.	16	1%	9	1%	3	1%	4	2%	
3d. Different survey approaches		Consider different survey mediums. Some staff prefer to stay with paper evaluations, some prefer online tools, recommend use of clicker technology. Surveys should be administered by an independent person, not the teacher as open to manipulation.	41	4%	26	4%	11	4%	4	2%	
4. Institutional use			Total Theme 4:	264	23%	163	24%	72	24%	29	17%
4a. Performance appraisal		Decrease importance/weight for promotion or don't use for this purpose. Change the way they are used in promotion process. Should be less compliance oriented. Not a suitable guide to teaching quality. Need more contextual information if data to be used by institution.	99	9%	73	11%	25	9%	1	1%	
4b. Abolish/negative impact		Negative impact on teachers, bureaucratic process. Current requirements too onerous. Abolish altogether. The data not reliable enough to be worth using.	25	2%	19	3%	3	1%	3	2%	
4c. Recognise limitations		Limitations/weaknesses should be recognised. Information is not very specific, an overview only.	18	2%	9	1%	6	2%	3	2%	
4d. Target towards specific staff only		Only of use to some staff. E.g. new/early career, where performance issues/low ratings.	5	0%	4	1%	1	0%	0	0%	
4e. Optional/decentralised		Should be optional and/or decentralised. Approach should be a departmental approach.	21	2%	17	2%	1	0%	3	2%	
4f. Compulsory/monitored		Should be compulsory and/or monitored. Need clear guidelines as to how often. Some staff suggested that made compulsory for students also. More coordinated approach. More cyclical with ongoing feedback.	49	4%	22	3%	14	5%	13	8%	
4g. Transparency, wider uses		Transparency of results to enable wider use by supervisors, HoDs, depts, students. Improve use for quality control. More use should be made of the results by the dept/university. More recognition of outstanding results.	30	3%	10	1%	16	5%	4	2%	
4h. Clarify purpose		Clarify purpose of evaluations, why used, recognise formal/informal functions and context. Separate the formative and summative purposes.	17	1%	9	1%	6	2%	2	1%	
5. No identifiable theme		Question not answered directly, meaning not clear enough to code or referred to previous question.	48	4%	30	4%	11	4%	7	4%	

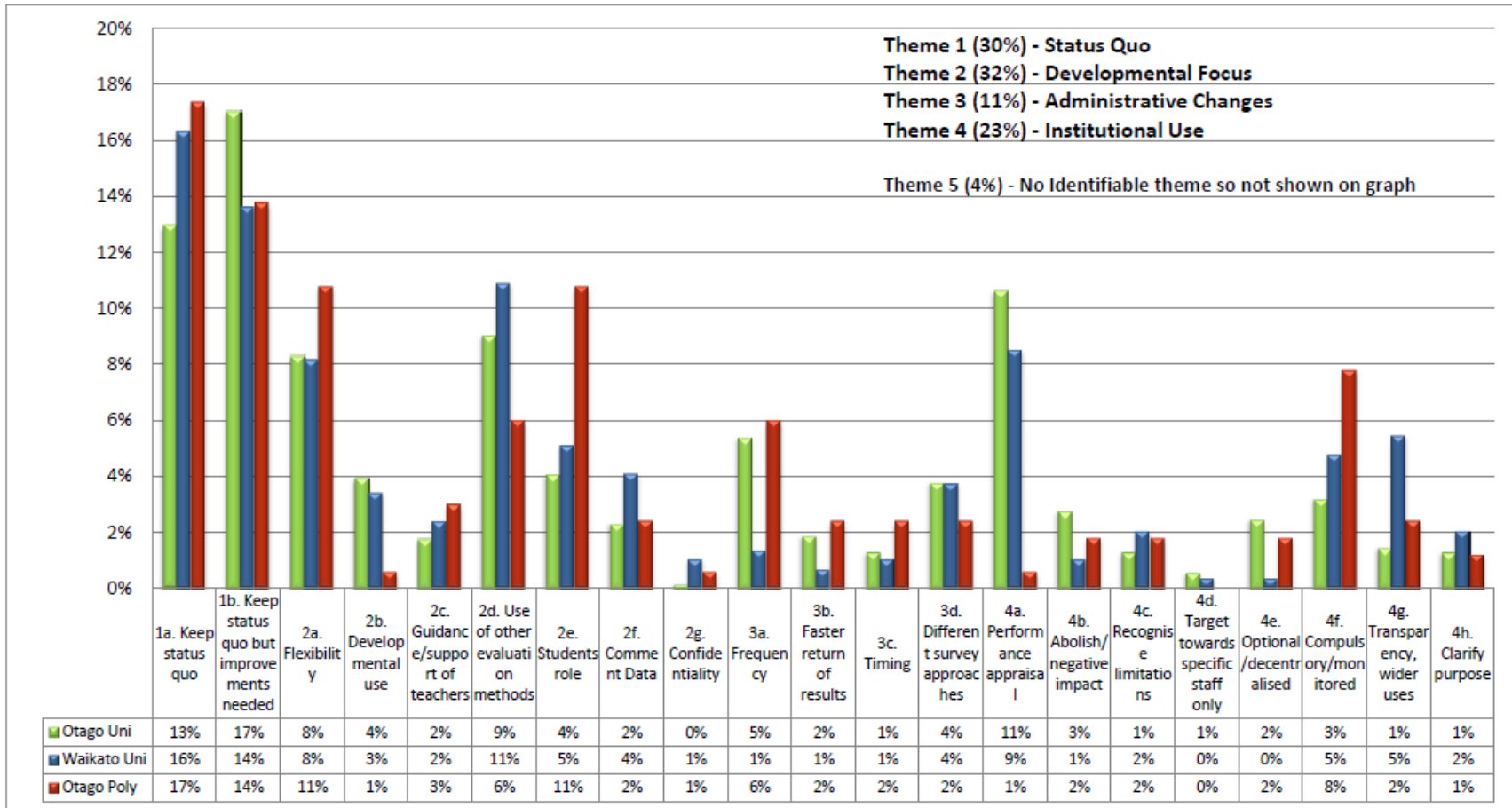
q21 Comment Data - Theme Coding - All Institutions

q21. If you were able to decide on the future of student evaluation/appraisal at your Institution, what would be your decision and why?



q21 Comment Data - Theme Coding - By Institution

q21 If you were able to decide on the future of student evaluation/appraisals at your Institution, what would be your decision and why?



Appendix 13 : Comment Data Analysis Q22

q22 If you have any general comments to make about your institution's appraisal process please make them here: (WU) / Do you have any other comments to make about student evaluation of teaching/courses? (OU and OP)

	Total Responses to	Num of q22 responses	%
Otago University (OU):	670	276	41%
Waikato University (WU):	242	102	42%
Otago Polytechnic (OP):	153	65	42%
Total:	1065	443	42%

Total All Theme Codes: 534

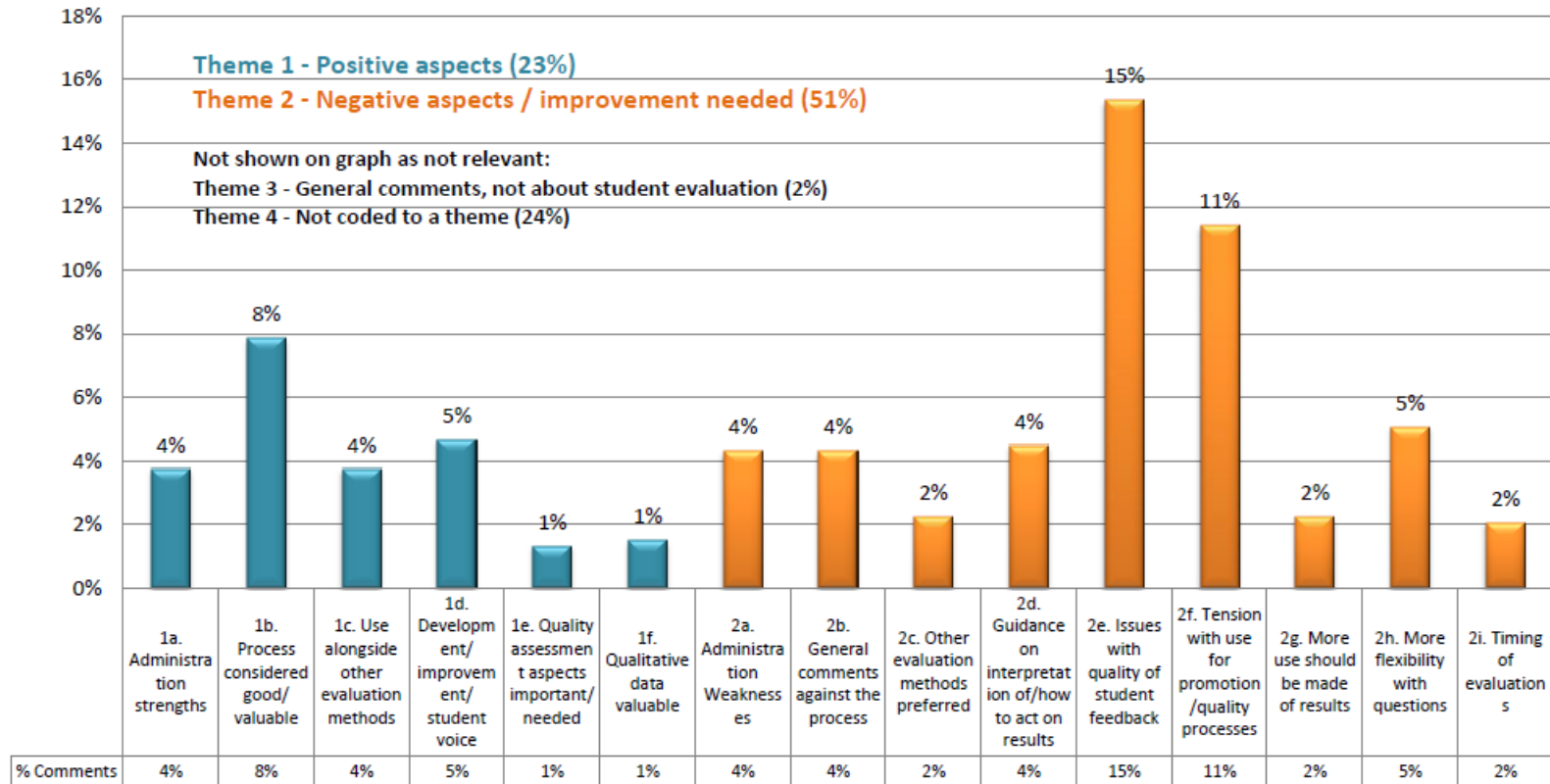
Theme & Sub-Theme Codes for q22	Additional Detail about each sub-theme	Total Count	Total %	OU Count	OU %	WU Count	WU %	OP Count	OP %
1. Positive aspects		Total Theme 1							
		122	23%	69	21%	35	26%	18	23%
1a. Administration strengths	Positive comments about administration aspects. Eg. turnaround times good, staff are helpful and provide good support. Efficient. Easy to use process.	20	4%	9	3%	9	7%	2	3%
1b. Process considered good/valuable	Generally happy with the student evaluation process. Value the process. Important that this is done. Useful/ valuable information is gained. Good basic method of data collection. Some benefits.	42	8%	20	6%	16	12%	6	8%
1c. Use alongside other evaluation methods	Better if used alongside other evaluation methods. E.g. class discussion, tutor feedback, outcomes, assessment results, peer/expert educator evaluation, externally facilitated focus groups.	20	4%	16	5%	1	1%	3	4%
1d. Development/ improvement/ student voice	Development: contributes to professional development of teacher. Helps inform improvement. Aids reflection. Student view: student voice is important. Good way to hear from a larger portion of the class than just a few.	25	5%	15	5%	4	3%	6	8%
1e. Quality assessment aspects important/ needed	Good to acknowledge good/excellent performance or improvement over time. Enables poor performance to be identified. Have a role to play in accessing quality but should emphasise the teacher's response to their evaluations rather than the results themselves. Quality aspects important but not if stifles innovation. Only really useful to early career teachers/ should be compulsory for new teachers.	7	1%	3	1%	3	2%	1	1%
1f. Qualitative data valuable	Comment data important. Should be able to routinely access comment data. Reasons why more important than the actual rating number.	8	1%	6	2%	2	2%	0	0%
2. Negative aspects/ improvement needed		Total Theme 2							
		275	51%	148	46%	78	59%	49	63%
2a. Administration Weaknesses	Negative comments about administration aspects. Eg. poor communication about what is expected especially for new staff, processing takes too long, need to be easier to complete, keep on 1-2 pages, process is time-consuming, format is boring (needs to be more dynamic and faster), unit is under-resourced, administration to students needs checking (to ensure carried out professionally). Look at providing them in different languages for non-English speaking students. Prefer online. Combine teacher and course evals into one. Teacher and course evals too similar.	23	4%	15	5%	4	3%	4	5%
2b. General comments against the process	Generally unhappy with process. Should not be done. Data not meaningful/useful/too broad. System is flawed, open to manipulation, cumbersome, inhibiting. Expensive when data is of such limited use. Better ways to improve teaching (co-mentoring and peer support).	23	4%	15	5%	5	4%	3	4%
2c. Other evaluation methods preferred	Other evaluations methods better/preferred, e.g. informal/custom surveys by lecturer/department, focus groups, informal discussions, tutor reports/feedback. Should give weight to course content, assessment and level.	12	2%	10	3%	1	1%	1	1%
2d. Guidance on interpretation of/how to act on results	How to interpret: caution needed as blunt instrument, need realistic expectations of results for large, compulsory, challenging/difficult papers, standardisation doesn't work well with all teachers/papers. Need benchmark/comparison data: compare like with like (e.g. level of course, class size, specialist vs service papers, would like to be able to compare with student assessments to be able to look at feedback from good and poor students), teaching context is important and should be viewed/presented alongside the appraisals (e.g. comparative workloads, amount of teaching involvement in the paper). Need better guidance on what to do with results, i.e. how to act on them and how to follow up with students.	24	4%	9	3%	10	8%	5	6%

q22 If you have any general comments to make about your institution's appraisal process please make them here: (WU) / Do you have any other comments to make about student evaluation of teaching/courses? (OU and OP)

Theme & Sub-Theme Codes for q22	Additional Detail about each sub-theme	Total Count	Total %	OU Count	OU %	WU Count	WU %	OP Count	OP %
2e. Issues with quality of student feedback	Issues with quality of student feedback/student ability to give constructive feedback. Students not capable of judging teacher quality/are not experts on pedagogy or content. Measure enjoyment/popularity only. Results can be skewed by lazy/disgruntled/malicious students (evaluations used to take 'revenge' on teachers). They don't take them seriously. They read the likert scale incorrectly. Students over-surveyed so make poor quality responses. Students need to be educated about the purpose of the process. These type of evaluations not suitable for all teaching, e.g. postgraduate/research supervision, small classes. Response rate concerns: move to online may reduce response rates too much, feedback given from students who don't attend class. If response rates too low reduces the value of the data (non-response bias & error rate).	82	15%	45	14%	15	11%	22	28%
2f. Tension with use for promotion/quality processes	Tension with use for promotion/quality processes. Use for promotions limits its formative use. Purpose should be clarified – summative or formative process. Separate these purposes. Not an accurate enough measure of teacher or course quality/value. Peer review of programmes and external assessment give better indications of quality. Can be penalised unfairly. Institution could be more supportive of staff and not penalise them too readily. Only seem to be used in promotion process. Should be voluntary not compulsory part of promotions process. Results should be confidential to the lecturer/do not make public.	61	11%	36	11%	21	16%	4	5%
2g. More use should be made of results	Little feedback about results eg from those assessing the value for PGS and promotion. Poor evaluations not acted on by management so no point and undermines the process for staff and students. More weight should be given to results especially in relation to research (view as equal). Confusion around who should have access to results. Results should be made public, at least within the institution. The process should be as transparent as possible. HoDs should get copy of results. Course coordinators should have access to results of teaching team evals. The evaluation process should be extended beyond just paper appraisal (programme, department, qualification level).	12	2%	1	0%	9	7%	2	3%
2h. More flexibility with questions	Required/core questions too limited, need reviewing. Questions need to be clearer. Separate standard from custom questions if part of promotion. Need flexibility with questions. Allow custom questions as well as core questions. More task-oriented questions. Need more questions on the effort students put into the course. Some questions are double-questions so not useful.	27	5%	15	5%	7	5%	5	6%
2i. Timing of evaluations	More guidance of use of evaluations other than at the end of semester (eg mid-semester). Timing important, shouldn't have to wait until the end of semester. Not possible to act on feedback as done too late in the semester. Do teacher evals during the semester and course ones at the end. Should be conducted later when students have had time to appreciate the learning (e.g. a year later).	11	2%	2	1%	6	5%	3	4%
3. General comments	Total Theme 3	10	2%	6	2%	4	3%	0	0%
3. General comments not specifically about student evaluations	Valuing teaching: the institution does not focus enough on teaching, teaching should be valued more, tertiary sector staff should all have teacher training, good to have something to bring the focus back to teaching rather than just research. General issues relating to the teaching environment: workload too high for spend much time on development, adds to stress of job with pressure on research output, out-moded teaching styles eg large lectures at expense of tutorials, lack of curriculum development within departments, staff being too individualistic and specialised instead of collaborative and working together to improve disciplinary knowledge, courses being 'dumbed-down' and students 'spoon-fed'. Comments about this research project: valuable and overdue, good to consider their usefulness.	10	2%	6	2%	4	3%	0	0%
4. Not coded to a theme	Total Theme 4	127	24%	100	31%	16	12%	11	14%
4. Not coded to a theme	Not able to be coded as: meaning unclear, referred back to previous question, haven't done any evaluations yet, answered in the negative (eg "No" or "No comment").	127	24%	100	31%	16	12%	11	14%

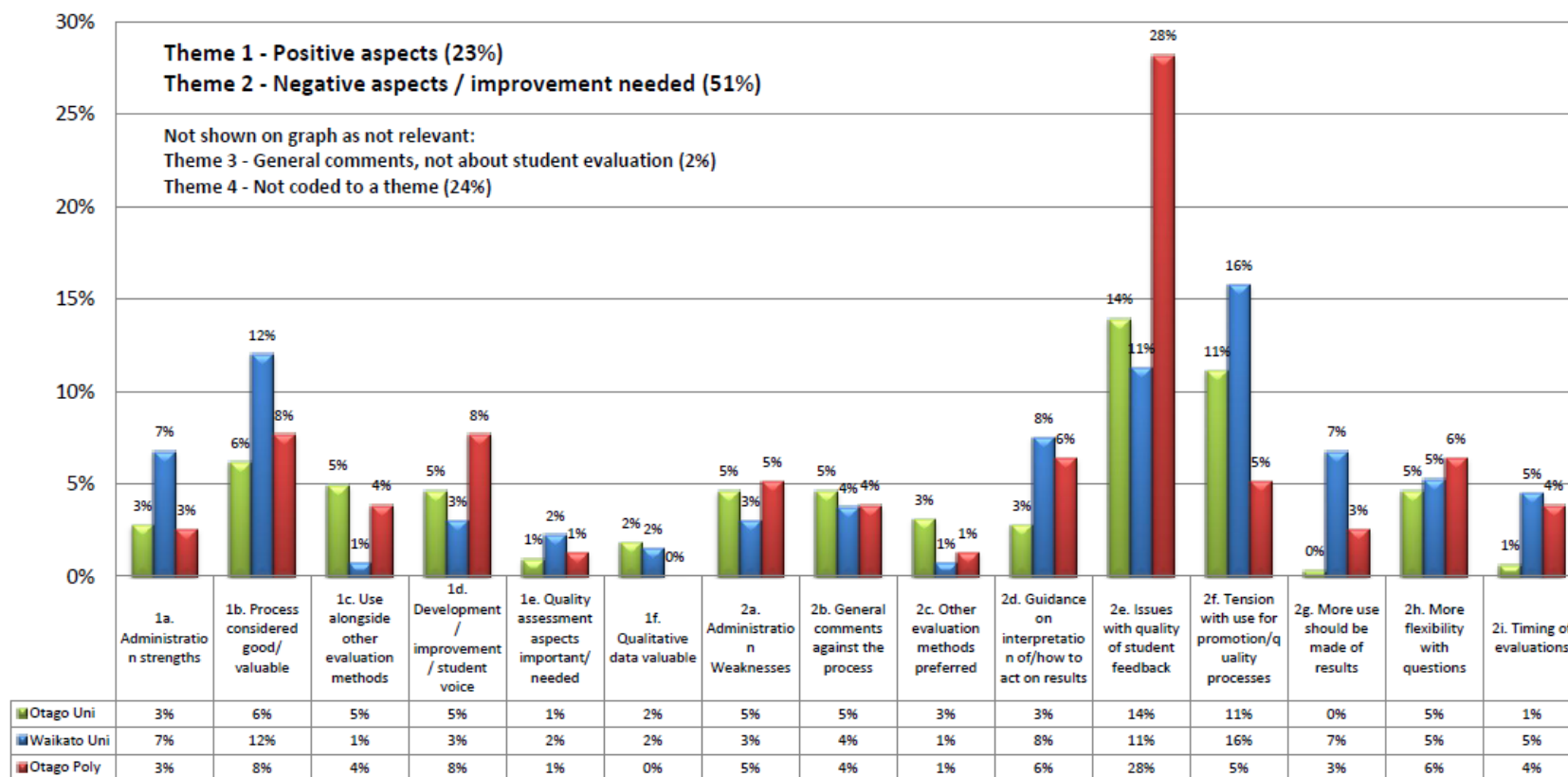
q22 Comment Data - Theme Coding - All Institutions

q22 If you have any general comments to make about your institutions appraisal process please make them here (WU) / Do you have any other comments to make about student evaluation of teaching/courses? (OU & OP)



q22 Comment Data - Theme Coding - By Institution

q22 If you have any general comments to make about your institutions appraisal process please make them here (WU) / Do you have any other comments to make about student evaluation of teaching/courses? (OU & OP)



Appendix 14 : Feeding Back to Students – An Example

Acknowledgment to Alison Campbell, University of Waikato

Email to students 2 December 2011

Hi everyone – I hope you're enjoying a well-earned summer break :-D

I've just received the 'extended answers' from the BIOL101 paper & teaching appraisals & I thought you might like some feedback on your comments & suggestions. (We don't get them back until after your grades have been confirmed, which is why I'm writing quite late.)

First up, thank you to everyone who took the time to comment on the teaching &/or the paper itself. We're always looking for ways to improve things so as to best support your learning, so it's always very helpful to hear from you. It's not always possible to act on every suggestion, though, & I'll try to explain why.

For example, quite a few people said that they'd like to get the 'lecture slides' in the study guide – I'm assuming that by this you're particularly keen on the various images. It's nice to know that you find those slides helpful! But, there are two reasons why I would not be keen to do this. One is that lecturers tend to change their lecture presentations every year – in my case, I might be doing this a few days before any given lecture. However, the study guides have to be with the Printery well before the semester begins... The other reason is that I really do want you to read more widely than the study guide! After all, this is essentially just a summary document. Much of what it contains is drawn from the textbook – and that's where most of the images we use in lectures come from. And yes, you really should be referring to the textbook on a reasonably regular basis. Plus, we have permission from the publishers to use their imagery in lectures, but *not* to include them holus-bolus in the study guide.

I was rapt that so many of you found my teaching style helpful to your learning – thank you for your comments, guys. This always means a lot to me but even more so this year as my first block of lectures (the molecular bio stuff) was 'new' to me as well as to you: I

had not given those lectures before & was very nervous about doing them well and in a way that let you get the most out of them :-) (Several people said I shouldn't change anything but I'm going to have to disagree – I believe that a teacher should always aim to improve!)

Quite a few of you noted that the lecture notes (in my section of the study guide) didn't always match up with what was covered in lecture. I'm sorry about that, & it is something that will definitely be fixed for next year. The reason it was like that this year is because I didn't know I'd be teaching that first block until quite late in the piece & had time only to write new lecture presentations; the study guide had already gone to Printery :-)

There were some comments in the 'should change' question, relating to the essay. I'm guessing these commenters would have preferred not to write one? :-) There are good reasons why we ask you to do this particular assessment item: much of the assessment in biology next year (& later) is going to require you to write essays, & I believe that giving you experience of that in first-year is going to help you perform better in subsequent years. Plus, one of the learning outcomes for the degree (it's even in the Graduate Profile) is that students will become competent in scientific communication, which includes writing – so again, we need to help you to learn how to do this well.

Tutorials... Having read through the comments, overall there were more in favour of MB than against. (Here I have to say, you're at university, people! I am *not* going to send out a weekly reminder to do tut work!) I can appreciate, though, that things like bandwidth can make them frustrating to do. The questions weren't always exactly on the lecture content, but that's part of the intention – to extend & deepen understanding.

Several people came up with what I think is a good compromise suggestion – that students could gain credit for either doing the MB tuts, or attending & *participating* in on-campus tuts - & I'll take this up with the rest of the teaching team. (Please note the word 'participating' – it would not be sufficient just to show up & sit there.) We had MB free this year, courtesy of the publisher (yay, Pearsons!), but next year students would have to buy access, so that compromise idea really does sound good. Plus there's obviously quite a strong preference for tutorials with a 'real' person :)

And it's nice to hear that the labs are so good for your learning & that those who used panopto, found it really helpful. I've learned from your comments that next year I need to do more to explain how & where to access the recordings, & also to show you how to download them (they're available as mp3 & mp4 recordings, but in retrospect that wasn't really obvious). I appreciate that you'd really really like it if everyone made use of this particular technology – but please do be aware that this is a personal choice for lecturing staff; everyone has a different teaching style & not everyone's comfortable with using panopto. So if you're keen on having it in a particular paper, do take time to approach the lecturer & explain why it's useful to your learning, but don't be upset if they in turn explain why they don't want to use panopto at this particular time.

Have a great couple of months & I'll look forward to seeing you again next year.

Alison

References

- Ako Aotearoa. (n.d.). *Ako Aotearoa New Zealand's National Centre for Tertiary Teaching Excellence*. Retrieved 8 August 2011, from <http://ako.aotearoa.ac.nz/ako-aotearoa>.
- Aleamoni, L. M. (1981). Student ratings of instruction. In J. Millman (Ed.), *Handbook of Teacher Evaluation* (pp. 110–145). Beverley Hills, Ca: Sage Publications.
- Arthur, L. (2009). From performativity to professionalism: Lecturers' responses to student feedback. *Teaching in Higher Education, 14*(4), 441–454.
- Ballantyne, R., Borthwick, J., & Packer, J. (2000). Beyond student evaluation of teaching: Identifying and addressing academic staff development needs. *Assessment and Evaluation in Higher Education, 25*(3), 221–236.
- Becher, T., & Trowler, P. (2001). *Academic tribes and territories. Intellectual enquiry and the culture of disciplines*. (2nd ed.). Buckingham: The Society for Research in Higher Education and Open University Press.
- Benton, S. L., & Cashin, W. E. (2012). *Student ratings of teaching: A summary of research and literature*. (IDEA Paper No. 50). Manhattan, KS: The IDEA Center. Retrieved 14 March 2012 from <http://www.theideacenter.org/category/helpful-resources/knowledge-base/idea-papers>.
- Benton, S. L., & Cashin, W. E. (2012). *Student ratings of teaching: A summary of research and literature*. (IDEA Paper No. 50). Manhattan, KS: The IDEA Center. Retrieved 14 March 2012 from <http://www.theideacenter.org/category/helpful-resources/knowledge-base/idea-papers>.
- Beran, T. N., & Rokosh, J. L. (2009). Instructors' perspectives on the utility of student ratings of instruction. *Instructional Science, 37*(2), 171–184.
- Beran, T. N., Violato, C., & Kline, D. (2007). What's the "use" of student ratings of instruction for administrators? One university's experience. *The Canadian Journal of Higher Education, 37*(1), 27–43.

- Biggs, J. B., & Tang, C. (2007). *Teaching for quality learning at university*. (3rd ed.). Maidenhead: Society for Research into Higher Education & Open University Press.
- Bovill, C. (2011). Sharing responsibility for learning through formative evaluation: moving to evaluation as learning. *Practice and Evidence of the Scholarship of Teaching and Learning in Higher Education*, 6(2). Retrieved 19 December, 2011 from <http://www.pestlhe.org.uk/index.php/pestlhe/issue/view/17>.
- Bowden, J., & Marton, F. (1998). *The university of learning: Beyond quality and competence in higher education*. London: Kogan Page.
- Boyer, E. (1990). *Scholarship reconsidered: Priorities for the professoriate*. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching, University of Princeton.
- Braskamp, L. A., & Ory, J. C. (1994). *Assessing faculty work. Enhancing individual and institutional performance*. San Francisco: Jossey-Bass.
- Burden, P. (2008). Does the use of end of semester evaluation forms represent teachers' views of teaching in a tertiary education context in Japan? *Teaching and Teacher Education*, 24, 1463–1475.
- Centra, J. A. (1993). *Reflective faculty evaluation*. San Francisco: Jossey-Bass.
- Costin, F., Greenough, W. T., & Menges, R. J. (1971). Student ratings of college teaching: Reliability, validity and usefulness. *Review of Educational Research*, 41(5), 511–535.
- D'Apollonia, S., & Abrami, P. (1997). Navigating student ratings of instructions. *American Psychologist*, 52(11), 1198–1208.
- Dye, J. F., Schatz, I. M., Rosenberg, B. A., & Coleman, S. T. (2000, January). Constant comparison method: A kaleidoscope of data. *The Qualitative Report* [Online serial], 4(1/2). Retrieved 29 September, 2011 from <http://www.nova.edu/ssss/QR/QR3-4/dye.html>.

- Edström, K. (2008). Doing course evaluation as if learning matters most. *Higher Education Research and Development*, 27(2), 95–106.
- Eisner, E. W. (1985). *The art of educational evaluation: a personal view*. London: Falmer Press.
- Eisner, E. W. (2002). From episteme to phronesis to artistry in the study and improvement of teaching. *Teaching and Teacher Education*, 18, 375–385.
- Eisner, E. W. (1998). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice*. Upper Saddle River, N.J.: Merrill.
- Erickson, F. (1998). Qualitative research methods for science education. In B. J. Fraser & K. G. Tobin, (Eds.) *International handbook of science education* (pp. 1155–1173). Dordrecht: Kluwer Academic Publishing.
- Fisher, R., & Miller, D. (2008). Responding to student expectations: a partnership approach to course evaluation. *Assessment and Evaluation in Higher Education*, 33(2), 191–202.
- Guba, E., & Lincoln, Y. (1989). *Fourth generation evaluation*. Beverley Hills, CA: Sage.
- Hendry, G. D., Lyon, P. M., Henderson-Smart, C. (2007). Teachers' approaches to teaching and responses to student evaluation in a problem-based medical program. *Assessment & Evaluation in Higher Education*, 32(2), 143–157.
- Ho, A., Watkins, D., & Kelly, M. (2001). The conceptual change approach to improving teaching and learning: An evaluation of a Hong Kong staff development programme. *Higher Education*, 42, 143–169.
- Kember, D., Leung, D., Y. P., & Kwan, K. P. (2002). Does the use of student feedback questionnaires improve the overall quality of teaching? *Assessment and Evaluation in Higher Education*, 27, 411–425.
- Laufgraben, J. L., & Tompkins, D. (2004). Pedagogy that builds community. In Laufgraben J. L., Shapiro, N. S., & Associates (Eds). *Sustaining and improving learning communities* (pp. 54–75), San Francisco: Jossey-Bass.

- Lucas, C. J. (1996). *Crisis in the academy: Rethinking higher education in America*. New York: St Martin's Press.
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11(3), 253–388.
- McKeachie, W. (1990). Research on college teaching. The historical background. *Journal of Educational Psychology*, 82(2), 189–200.
- McKeachie, W. (1997). Student ratings: The validity of use. *American Psychologist*, 52(11), 1218–1225.
- Menges, R. J., & Mathis, B. C. (1988). *Key resources on teaching, learning, curriculum and faculty development*. San Francisco: Jossey-Bass.
- Miller, S., & Fredericks, M. (2003). The nature of "evidence" in qualitative research methods. *International Journal of Qualitative Methods*, 2(1). Article 4. Retrieved 26 October 2011 from <http://www.ualberta.ca/~ijqm>.
- Millman, J. (Ed.). (1981). *The handbook of teacher evaluation*. Beverley Hills CA: Sage.
- Ministry of Education. (2010). *Tertiary Education Strategy 2010–2015*. Retrieved 18 October 2011 from <http://www.tec.govt.nz/Tertiary-Sector/Tertiary-Education-Strategy/>.
- Mintzberg, H. (2004). *Managers not MBAs: A hard look at the soft practice of managing and management development*. San Francisco: Berrett-Koehler Publishers.
- Moore, S., & Kuol, N. (2005). Students evaluating teachers: Exploring the importance of faculty reaction to feedback on teaching. *Teaching in Higher Education*, 10(1), 57–73.
- Nasser, F., & Fresko, B. (2002). Faculty views of student evaluation of college teaching. *Assessment and Evaluation in Higher Education*, 27(2), 187–198.
- Ory, J. C., & Ryan, K. (2001). How do student ratings measure up to a new validity framework? *New Directions for Institutional Research*, 27–44.

- Otago Polytechnic. (2011). *Otago Polytechnic strategic directions 2011–2013*. Retrieved 20 September, 2011 from http://www.otagopolytechnic.ac.nz/fileadmin/DepartmentalResources/Marketing/Strategy_2011-2013/OP-Strategic_Directions_2011_-_13_final.pdf.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage Publications.
- Patton, M. Q. (1997). *Utilization-focussed evaluation: The new century text*. Thousand Oaks, CA: Sage Publications
- Penny, A., & Coe, R. (2004). Effectiveness of consultation on student ratings feedback: A meta-analysis. *Review of Educational Research*, 74(2), 215–253.
- Pratt, D. D. (1992). Conceptions of teaching. *Adult Education Quarterly*, 42(4), 203–220.
- Prosser, M., & Trigwell, K. (1999). *Understanding learning and teaching: The experience in higher education*. Buckingham: SRHE and Open University Press.
- Ramsden, P. (1992). *Learning to teach in higher education*. London: Routledge.
- Ramsden, P., & Dodds, A. (1989). *Improving teaching and courses: A guide to evaluation*. Melbourne: University of Melbourne Centre for the Study of Higher Education.
- Schmelkin, L. P., Spencer, K. J., & Gellman, E. S. (1997). Faculty perspectives on course and teacher evaluations. *Research in Higher Education*, 38(5), 575–592.
- Schwandt, T. A. (1994). Constructivist, interpretivist approaches to human inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 118–137). Thousand Oaks, California: Sage Publications.
- Schwandt, T. A. (2007). Judging interpretations. *New Directions for Evaluation*, 114, 11–14.
- Silverman, D. (2001). *Interpreting qualitative data*. (2nd ed.) London: Sage.
- Smith, C. (2008). Building effectiveness in teaching through targeted evaluation and response: Connecting evaluation to teaching improvement in higher education. *Assessment and Evaluation in Higher Education*, 33(5), 517–533.

- Smith, M. K. (2005). Elliot W. Eisner, connoisseurship, criticism and the art of education. *The encyclopaedia of informal education*. Online at www.infed.org/thinkers/eisner.htm. Accessed 19 December, 2011.
- Smock, H. R., & Crooks, T. J. (1973). A plan for the comprehensive evaluation of college teaching. *Journal of Higher Education*, 44, 577–586.
- Stronge, J. (2006). Teacher evaluation and school improvement. In J. Stronge (Ed.), *Evaluating teaching: A guide to current thinking and best practice* (pp. 1–23). Thousand Oaks, CA: Corwin Press.
- Theall, M., & Franklin, J. (2001). Looking for bias in all the wrong places: A search for truth or a witch hunt in student ratings of instruction? *New Directions for Institutional Research*, 109, 45–56.
- Trigwell, K., & Prosser, M. (1996a). Congruence between intention and strategy in university science teachers' approaches to teaching. *Higher Education*, 32(77–87).
- Trigwell, K., & Prosser, M. (1996b). Changing approaches to teaching: A relational perspective. *Studies in Higher Education*, 21(3), 275–284.
- The University of Waikato. (2009). *University strategy 2010–2013*. Retrieved 20 September, 2011 from <http://www.waikato.ac.nz/about/corporate/strategy.shtml>.
- University of Otago. (2005). *The teaching and learning plan 2005–2010*. Dunedin: University of Otago. Retrieved 18 January, 2012 from <http://hedc.otago.ac.nz/tlp/home.do>.
- University of Otago. (2003). *University of Otago charter*. Retrieved 20 September 2011 from <http://www.otago.ac.nz/about/otago000764.pdf>.
- University of Otago. (2011). *2010 Annual report*. Dunedin: University of Otago. Retrieved 1 September, 2011 from http://www.otago.ac.nz/about/official_documents.html#annualreport.
- University of Otago Human Resources. (2011). *Academic staff promotions. Guidelines for applicant, Heads of Departments/Deans, administrators and promotion committees*. Retrieved 17 October 2011 from

<https://docushare.otago.ac.nz/docushare/dsweb/Get/Document-11235/AcademicPromotionsPolicy.pdf>.