



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Martinez Marco, Lucia Alicia

Title:

Metaevaluation in Australian University Language Programs

Date:

2020

Persistent Link:

<https://hdl.handle.net/11343/242501>

Terms and Conditions:

Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.

Metaevaluation in Australian University Language Programs

Lucía Alicia Martínez Marco

ORCID

0000-0001-9195-1888

May 2020

Submitted in total fulfilment of the requirements of
the degree of Doctor of Philosophy

Abstract

Calls to conduct research on evaluation have been made to learn about its practical realities, in order to ensure a quality process and to improve future evaluation experiences. A key methodology addressing this need is metaevaluation, the evaluation of an evaluation. Studying seven Australian university case studies of reviews of language programs, meta-evaluation is used as both the conceptual framework as well as a methodological tool to explore and assess the evaluation practices and to discover practice patterns, using the findings to make recommendations about how such evaluations should be conducted.

The research in this thesis demonstrates that New Public Management (NPM) which operates as a common organizational and policy framework in Australian Higher Education and has been defined as a national and local control system which “models higher education simply in economic terms” (Marginson 2010, p. 6) is the most influential evaluation factor in reviews.

In this policy context, a gap is found between the tenets of evaluation theory and the character of Australian Language Program Evaluations (aligned with NPM ideology). As a consequence, most initiators or sponsors of evaluations are not focused so much on the quality of the process but on using reviews as a top-down instrument to implement a cost-efficiency agenda.

As the analysis of data progressed, it became clear that the majority of cases reveal a process of reviewing empty of any theoretical foundations based in the evaluation discipline, thus putting in jeopardy the quality of the evaluation process, which resembled more a sponsors’ top-down exercise of checking or imposing compliance with University priorities. In light of these broad findings, this thesis claims that metaevaluation has a key role in raising the quality of the practice of reviewing by ensuring reviews are informed by evaluation theory and practice, thus following principles of good practice, and that it can also help mitigate a purely cost-driven process.

Key Words

university foreign language programs; program evaluation; metaevaluation; language program reviews; university language teaching; Australian language program evaluation; foreign language learning; neoliberal education policy; New Public Management; higher education policy; higher education management; stakeholders

Acknowledgments

Thanks for all their loving care and support to my family, particularly my husband Vicente, my son Marco and my mother Marita; to my friends, especially Robe and Mercedes; to all my NHS Scottish doctors and nurses.

My deep gratitude to Prof. Tim McNamara, my academic and life mentor, unique supervisor, friend and scholar; to all the dissertation panel members lead by Tim, particularly Carsten; to all my great teachers in my Honours and PhD courses at the University of Melbourne, and in my MA course at Indiana University; to LCNAU and its supervising committee, led by Prof. Hajek and Dr. Roever.

Thanks for their support to the University of Melbourne, in which School of Languages and Linguistics my research ideas have been enthusiastically nurtured and I have substantially progressed in my Applied Linguistics academic career; to all collaborating Australian institutions for facilitating essential documents for this thesis; to all participants who have agreed to cooperate in this project by offering their honest opinions and sharing their academic experiences.

This dissertation is dedicated to all of you! Metathanks!

Table of Contents

List of Figures	viii
List of Tables	viii
List of abbreviations	viii
1. Literature Review	
1.0 Prologue	3
1.1. From Evaluation to Metaevaluation	4
1.1.1. Metaevaluation: The Latest Advance in Evaluation Theory to Improve Evaluation Practice and Evaluands	13
1.1.2. Metaevaluative process challenges	17
1.1.3. From <i>Doing</i> to <i>Using</i> Evaluation and Metaevaluation as Instruments for Language Program Improvement	20
1.1.4. Using Language Program Metaevaluation	23
1.1.5. Fostering an on-going Utilization-Focussed LPE evaluation	25
1.1.6. Fostering an Ongoing Evaluation in LPE	32
1.1.7. Nevo's Dialog Evaluation Approach (Nevo 2009)	35
1.1.8. From Evaluation to Metaevaluation: Summary	36
1.2. Context and Metaevaluation	38
1.2.1. Metaevaluating Context 1: Language Program Evaluation and Metaevaluation in Australia	40
1.2.2. Metaevaluating Context 2: Historical Developments in Evaluation in Australia	42
1.2.2.1. Process of Professionalization of Evaluation in Australia	43
1.2.2.2. Evaluation for Accountability versus Evaluation for Improvement in a Managerial Culture	46
1.2.3. Metaevaluating Context 3: Understanding the University New Public Management Business Model Built Upon Neoliberal Ideology	49
1.2.3.1. Embracing NPM Principles to Overcome a "Malaise"	50
1.2.3.2. The Impact of the NPM Policy Context on Educational Quality	58
1.2.3.3. Summary and Research Questions	61

2. Conducting Metaevaluation	65
2.1. Introduction	65
2.2. A Multiple Research Methodology Approach for Metaevaluation	65
2.2.1. The Measuring Instruments: Metaevaluation Criteria and Metaevaluation Quality	67
2.2.2. The Metaevaluator	70
2.2.3. Metaevaluation Focus	73
2.3. An Emergent Realist Evaluation (ERE) Model for Conducting Metaevaluation	74
2.3.1. A Taxonomy of Subjects and Inquiry Modes for Metaevaluating LPEs in Australian HE	78
2.3.1.1. Purpose(s) of Metaevaluation	79
2.3.1.2. The Description Mode	83
2.3.1.3. The Classification Mode	84
2.3.1.4. Values-Inquiry	84
2.3.1.5. Taxonomy of Subjects of Inquiry	85
2.4. Metaevaluation Data Analysis: Evidence, Collection & Methodologies	86
2.4.1. Review Reports	88
2.4.2. Interviews	92
2.4.3. Methodologies for Data Analysis	94
2.4.3.1. Transcription of Interview Data	94
2.4.3.2. The Case Study (CS)	97
2.4.3.3. Document Analysis (DA)	99
2.4.3.4. Thematic Analysis (TA)	100
2.4.3.4.1. Coding for TA	103
2.5. Connection of Analytical Framework With Overall Metaevaluation Framework	104
2.6. The Role of Context in Metaevaluation and Implications for this	107
3. Findings	110
3.1. Executive Summary of Report's Findings	110
3.1.1. Program description	111
3.1.2. Brief description of methodologies and analytical strategy	112
3.1.3. Evaluation questions and evaluation purpose(s)	113
3.1.4. Areas of study	113
3.2. Aligning the Languages Schools and Programs with University Priorities. Context	

in CS2, CS4 and CS5	116
3.2.1. Evaluation questions and evaluation purpose in CS1, CS6, CS3, and CS7	118
3.2.2. Recommendations	119
3.3. The Table of Contents (ToC) and Other Sections that Preface the Report	121
3.3.1. Description	121
3.3.2. Findings	121
3.3.3. List of Tables, Graphs and/or Figures; Acronyms and Abbreviations; Acknowledgements	121
3.3.4. Additional Section: The Terms of Reference.	123
3.3.5. Terms of Reference: Membership of the Review Panel (RP).	135
3.4. Evaluand and evaluation context	146
3.4.1. Description	146
3.4.2. Evaluand, Program Goals and Objectives	147
3.4.3. Evaluation Audiences and Target Population	148
3.4.4. Review of Related Research	155
3.4.5. The Structure of the Report	157
3.5. The Methodology Chapter	158
3.5.1. Description	158
3.5.2. Findings	159
3.6. The Results Chapter	178
3.6.1. Description	178
3.6.2. Findings	178
3.6.2.1. CS1	179
3.6.2.2. CS6	181
3.6.2.3. CS3	184
3.6.2.4. CS7	190
3.7. Summary, Conclusions, and Recommendations	194
3.7.1. Description	194
3.7.1.1. Summary of Findings	195
3.7.2. Conclusions and Recommendations	203
3.7.2.1. CS2, CS4, and CS5	203
3.7.2.1.1 Conclusions	203
3.7.2.1.2 Recommendations in CS2, CS4 and CS5	203
3.7.3. CS1, CS3, CS6 and CS7	207

3.7.3.1. C6	207
3.7.3.2. CS3	211
3.7.3.3 CS1	212
3.7.3.4. CS7	215
3.8. References and Appendices	217
3.8.1. Description	217
3.8.2. Findings	217
3.8.3. Appendices: Clearly Presented and Explained; Cited in Text; Attached in the Same Order as Cited	219
3.8.3.1. CS3	222
3.8.3.2. CS4	222
3.8.3.3. CS6	227
3.8.3.4. CS7	228
4. Discussion	230
4.1. Introduction	230
4.2. Strategic Policy Use of Quality Assurance Evaluation for Accountability in Australian HE Preventing Improvement	232
4.2.1 Academics vs. Managers	245
4.2.2. Managerial agenda implementation dominating the evaluative process	246
5. Recommendations and Conclusion	250
5.1. Recommendations	250
5.1.1. Utilization-Focussed LPE evaluation	250
5.1.2. Limitations of the study: Criteria for metaevaluation	254
5.2 Conclusion	255
Works Cited	256
Endnotes	269
Appendices	
Appendix 1. Interview Questions	277
Appendix 2. Conventions for the transcription of interviews' verbal data	279
Appendix 3. Letters to Institutions	280

Appendix 4. Case studies EFI Tables	282
Appendix 5. Case studies coding	299

List of Figures

Figure 1: Interviewees	93
Figure 2: Multiple methodologies for a metaevaluation of Australian Language Program Evaluations (ALPE)	97
Figure 3: Linking multiple methods	104

List of Tables

Table 1: Some language program evaluations in Higher Education in the United States	29
Table 2: Taxonomy of subjects for a metaevaluation of ALPE	106
Table 2: Inquiry Modes and Thematic Analysis	107
Table 3: Structure of Report and Recommendations, CS5	120
Table 4: The range of interview participants and languages represented in CS2	164
Table 5: Areas under review, CS2 and CS3	184
Table 6: Structure of Report, CS3	186
Table 7: Range of interview participants and languages represented in CS4	225

List of Abbreviations

CS1, 2, etc.	Case Study
TA	Thematic Analysis
EFIS	Evaluation Fundamental Issues
DA	Document Analysis
ERC	Evaluation Report Checklist
RP	Review Panel
ES	Executive Summary
ToC	Table of Contents
ToR	Terms of Reference
A1, A2, etc.	Appendices
HoS	Head of School
UG	Undergraduate Program

Metaevaluation in Australian University Language Programs

Chapter 1. Literature Review

Chapter 1. Literature Review

1.0 Prologue¹

This thesis is a metaevaluation of seven example of program reviews in languages departments or Schools in Australian universities. The term ‘program reviews’ is commonly used in Australia to refer to what elsewhere might be called program evaluations. The examples will be referred to throughout the thesis as ‘case studies’. The argument of the thesis is supported by two kinds of data. The first and principal data source consists of review reports. The findings of a systematic analysis of these reports is supplemented by data from interviews with a number of individuals who have been involved in review processes either as reviewers or as those involved in the delivery of language programs, who can be seen as being at the receiving end of the reviews. The interview data is not presented systematically in the same way as in the reporting of the documentary analysis, but is used to flesh out and contextualize features of the reports.

Languages departments in Australian universities often don’t have program evaluation as an embedded component in their regular programmatic activities (Martínez Marco 2013). So, how do they monitor their practices? What character do program reviews conducted in language departments have? Who plans them? Who conducts them? For what purpose? What methodologies are used? Are these practices informed by Evaluation theory and practice? How can we learn the answers to these questions?

In the field of evaluation, calls to conduct research on evaluation have been made to learn about the practical realities of evaluation, in order to ensure a quality process and to improve future evaluation experiences. This impetus is believed to contribute to advancing

the evaluation profession. One of the concepts and methodologies that have been claimed to address this need is metaevaluation, that is, the evaluation of an evaluation. Using seven Australian universities case studies, metaevaluation is used in this thesis as both the conceptual framework as well as a methodological tool to explore evaluation practice, to assess the quality of evaluation practice in its application, to discover potential practice patterns and to use the findings from the metaevaluation to make recommendations.

Judging from the still exploratory research on evaluation in higher education and particularly of teaching, which is still “a field under construction” (Stake, Contreras and Arbesú 2012, p. 11), the internal organization of a university is important to the efficiency of its operations, and evaluation is highlighted as ‘a good helper’ for administrators to making informed decisions (Kogan 1989, Secolsky and Denison 2012). But what are the policies that govern Australian universities’ activities? What is their organizational framework? These are some of the issues which this thesis will explore.

1.1. From Evaluation to Metaevaluation

What exactly is evaluation and how is it defined? In essence, “evaluation is part of everyone’s normal behavior” (Stake, Contreras and Arbesú 2012, p. 4) and as such, it refers to the value judgments in which all humans engage day to day when they interact with the world around them. This kind of activity usually takes place in an unconscious and informal manner and it manifests in everyday life activities such as when valuing the food we eat, the type of clothes to wear for a job interview, etc. (Owen 1999). These judgments feed into the structure of meanings (sense-making) supported by each individual's idiosyncratic values and beliefs and are embedded in the ongoing chain of decisions we all make.

Evaluation as a discipline was conceived in the US in the 1960s in an era when “social

progress has become the public business” (Cronbach and associates 1980, p. 12).

Furthermore, Program Evaluation (PE) is a distinctive professional practice which resulted from the lessons learnt from the evaluations during the economic depression of the 1930s in the US, which slowly led to the Great Society of the 1960s when the government injected substantial federal money for improving social services, including education (Patton 1997).

The passage of the Elementary and Secondary Education Act of 1965 (Smith 2009) was a stepping-stone for introducing evaluation for course improvement in a systematic way (Cronbach 1963). Many evaluations of educational programs during the 1970s and a good part of the 1980s were funded with public money and commissioned by the US government, mainly for curriculum improvement in secondary education. In fact, the beginnings of educational evaluation are considered a “side effect of curriculum reform” resulting from an intended revision of what was taught in the science and mathematics secondary education class by the National Science Foundation after the Soviet Union satellite, Sputnik, was launched in 1957 (Cronbach et al 1980, p. 31). Therefore, the origins of evaluation, particularly of program evaluation, are closely tied to raising educational standards for the improvement of education.

It has been recognized that while the capacity to evaluate is natural and basic to all humans, it lacks the formal and systematic thinking that is needed for more thorough and defined evaluative tasks, which is what became known as *formal evaluation*. This is not to say that formal evaluation should replace natural sensemaking; only that evaluation is an *assisted technology*, as the term indicates, to aid it (Mark, Henry and Julnes 2000).

Therefore, informal evaluation should not be disregarded. As Schwandt and Burgon (2006) argue: “Everyday, informal evaluation takes place without them [evaluators], that their systematic work occurs in a world filled with natural evaluative judgments, and that it

is but a small fraction of that larger evaluative world” (p. 3). More recently, informal evaluation has been characterized as an expression of a reflexive culture, that is, examples of embedded evaluation moments (McClusky et al 2008). As an illustration, it is acknowledged that teachers engage in evaluation activities “though often informally” (Darling-Hammond 1990, p. 141) in educational contexts. However, some theorists have come to distinguish these informal reflective exercises from actual *evaluative practice* defined as “observable, practical situations” (Saunders 2011, p. 15) thus emphasizing the importance of aiding informal evaluation with formal evaluation. But as a formal assisted technology and rooted in academic work, can evaluation be distinguished from ‘research’?

Evaluation theorists such as Alkin (1975) support the practical nature of evaluation versus knowledge-production, which is the object of research. Thus, a key dimension that helps distinguish evaluation from other forms of research is evaluation *utilization* (Patton 1978), that is, using it to shape educational programs. At the same time, however, it is one of the issues which keeps emerging as problematic, because evaluators often have to face the ‘non-utilization’ of findings after the evaluation has been conducted and the issue remains unresolved today as to what makes utilization *click* (Shaw, Greene and Mark 2006). Increasingly, though, the concept of utilization has been broadened by including indirect uses of evaluation (Patton 1997, 1998, 2002).

By conducting evaluation informed by evaluation principles and metaevaluation, there is a more realistic chance of utilization, that is, *using* is emphasized over *doing* evaluation. There is also a greater chance of benefiting from *process use*, that is of having “changes in thinking, attitudes and behavior, and program and organizational changes in procedures and culture” (Patton 2008, p. 155). This suggests that the knowledge acquired in the process of conducting an evaluation informed by evaluation principles can teach us not only about

the procedures involved in conducting an evaluation (i.e. what to pay attention to in an evaluation), but more importantly will teach evaluative thinking defined as “a logical mode of thought informed by systematic, rational, and purposeful information gathering. The application of evaluative thinking results in organizational decision-making being driven by reasoned choices based on empirical evidence” (Davis, Sinicrope and Watanabe 2009, p. 216). In this context, evaluation participants will be engaged in an ‘evaluation culture’ which has a longer life-span than findings themselves (Patton 2002). In sum, there is no doubt that many are the potential benefits, both direct and indirect, short and long term, of the formal application of evaluation as an assisted technology to help natural sense-making.

Evaluation as a *transdiscipline* (Scriven 2009) presents a history labeled as ‘complex’ by various authors at different times. It reflects a number of challenges besides the above-mentioned use of the concept of utilization. There has been difficulty in reaching a consensus in defining evaluation formally. Thus, different definitions have been suggested depending on the area of focus: (a) the merit or worth of a program; (b) the function of evaluation as social research; (c) the political dimension of evaluation; and (d) the distinction between monitoring and evaluation (Wilson and Mertens 2012). The acquisition of an understanding of evaluation frameworks, to inform evaluation practice is thus key to help to inform decisions on how to go about improving Language programs. In the particular realm of education, a uniform definition has been virtually unattainable due to (1) the diverse kinds of evaluands (described by Scriven as the object under evaluation (1981)); (2) the different kinds of localities where evaluands are found; (3) the variety of social perspectives with which evaluation practice may be assessed; (4) the range of methodologies guided by different schools of thought; (5) the wide array of discourses that evaluation draws on; and (6) the lack of credentialed status of evaluation in comparison to

other disciplines such as medicine or law (Schwandt 2009). Nevertheless, among the diversity of definitions found in the literature, Trochim suggest a moderately comprehensive view of evaluation:

[Evaluation] is a profession that uses formal methodologies to provide useful empirical evidence about public entities (such as programs, products, performance) in decision-making contexts that are inherently political and involve multiple often-conflicting stakeholders, where resources are seldom sufficient, and where time-pressures are salient (Trochim 1998, p. 248).

The value of this definition lies in that it reveals some of evaluation's most fundamental issues such as the purposes of evaluation, formal evaluation methodologies and evaluands, as well as key evaluation dimensions such as the political one and other realities which involve practical constraints that challenge the practice of evaluation. Smith (2009) highlights the relevance of these dimensions by stating that "as the world changes, the answers to these underlying issues will also continue to change; thus, any contemporary answer to a fundamental issue is only a temporary resolution. The fundamental issue, however, remains" (p. 48-49), which in turn implies the tight relationship between evaluation and its context.

The eminent evaluator Cronbach (1982), while explaining the purpose for which his book *Designing Evaluations for Social and Educational Programs* was written, highlights the importance of being aware of what it implies to evaluate: "I seek to sensitize the prospective evaluator and the critic of evaluation as to the innumerable decisions that should lie behind what the evaluator does" (p. ix). More recently and in the same vein, Simmons (2006) emphasizes the external influences which impinge upon the ethical dilemmas that evaluators face when making decisions about different courses of action depending on "a myriad of factors - social, personal, political, cultural - that are pertinent in the particular context" (p. 243).

Beyond the scope of the evaluators' multiple decisions, further external as well as internal arguments contribute to define the complexity of evaluation. First, there is the "complex contextual landscape of evaluation as a profession, occupation, community, and discipline" (Mark, Greene and Shaw 2006, p. 18). Second, there is the mix of traditions that have influenced evaluation (Mertens and Wilson 2012), particularly, the wide variety of academic backgrounds of those involved in evaluation, each of which "has had its own tradition in evaluation, its own criteria of proper procedure and its own respected sources" (Cronbach et al 1980, p. 44). Finally, as an example of evaluation's internal intricacies, Dahler Larsen (2009) mentions the "collective, contextual, relational, organizational and political" traits that characterize the very nature of the evaluation process. This variety of influences is reflected in the diverse views expressed by different groups of evaluation theorists/evaluators about all evaluation aspects, from, as seen above, defining evaluation to ways to approaching fundamental evaluation issues, which have driven evaluation theory to ensure that the evaluation process is adequate, formal and ethical, addressing and overcoming challenges that emerge in an evaluation.

While the conceptual foundation represents one of the most important dimensions in evaluation and has offered ample substance for advancing evaluation's theoretical debate, it has mainly contributed to the intellectual aspect of the discipline. On the whole, what the evaluation literature shows in its evolutionary stages are ongoing attempts to legitimize the worthwhileness of evaluation. Hence, many proposals for *theories of evaluation* have been formulated, and many issues have been brought up for discussion seeking to refine the concept and practice of evaluation. In particular, different epistemological and ontological perspectives associated with specific methodological procedures have sparked heated intellectual discussions motivated mainly by basic questions such as 'how is it done?' and

‘what works?’ without arriving at a consensus as to what constitutes good practice in evaluation (Norris 2008).

Thus, the methodologies suggested for conducting evaluation have been developed over time from theories of evaluation reflecting different world views and representing one of the most important debates in evaluation known as the paradigm debate. The *positivistic* school of thought, associated with quantitative methodologies, and the *naturalistic*, associated with qualitative ones, constitute the philosophical source from which theories’ tenets are formed and characterized. The rationalistic view of studying a program with a prefixed plan with pre-specified questions dominated evaluation practice through the 1970s, to the extent that it reflected the neglect of the proponents of this movement to open the discussion to consider broadening the conceptualization of evaluation. However, this view of evaluation posed central challenges such as the knowledge-production orientation relegating its instrumental utility, which have led evaluators to emphasize evaluation’s decision-making function. Focusing evaluation on the estimate of net effects by addressing a particular question by a decision-maker was also criticized since decision-making should entail a process involving a plural community as in Cronbach’s *policy shaping community*, (1982), which depicts social programs as a site of multiple competing perspectives in which the evaluator’s main role would be to negotiate a consensus.

Therefore, when evaluators address, respond to and serve the needs of those who commission the evaluations more than “the issues of moment to other actors in the program context” (Weiss 1986, p. 147), questions of which interests should evaluation serve, and also, of stakeholding are raised². In recent times, this key political dimension of evaluation has revived in the light of the increasing accountability movement in education, particularly in universities and therefore pertinent to consider in this thesis, where the risk

of bias is visible since

only the intentions from management are used to judge the worth of a program or practice. Participants are solely taken serious for the information they are able to provide. They are approached instrumentally. This does not lead to acceptance and implementation of evaluation data and results (Niessen et al 2009, p. 377-378).

Therefore, stakeholder involvement implies being part of the evaluation team and participating in “planning and framing; data collection, analysis and interpretation; and reporting and use” (Ryan and Cousins 2009, p. 551) rather than simply constituting a resource for data collection.

Hence, as evaluators apply methods in real evaluation experiences, as new dimensions emerge and are added to the discussion and traditional ones continue to be explored, advocates of evaluation pursue the search for arguments to legitimize evaluation under a particular perspective. But regardless of the complexities of evaluation, what remains clear is the existence of a body of knowledge, theoretical paths and methodological models to conduct evaluations *formally* to ensure a quality evaluation process. This source of knowledge also makes us aware of important challenges for evaluation practice that cannot be ignored which need to be taken into consideration when planning the evaluation.

Therefore, the application of evaluation theory in evaluation practice is key for any reputed institution, particularly in an educational setting, as in this thesis, seeking to conducting evaluations that can help to make informed decisions for improvement. Otherwise, the validity of findings, a key concept in evaluation, and also in assessment and measurement (Secolsky and Denison 2012a) can be put into question, thus affecting evaluation utilization as well as bringing about a variety of further, foreseeable or not, negative consequences. Furthermore, educational institutions can be at a disadvantage since conducting evaluations that ignore evaluation tenets can be perceived as a lack of commitment to good practice.

Validity is associated with how value judgments are arrived at and how the conclusions derived from those judgments are expressed since they will influence the credibility of findings. When evaluating a program formally, “because validity is linked to the meaning, value and appropriateness of interpretation, validity is the most critical consideration in research and evaluation” (Cisneros-Cohernour 2012, p. 511). To address validity in evaluation, evaluators such as Stake (2004) and Patton (1997) claim the use of triangulation and so does Grayson (2012) in the evaluation of programs in higher education. Triangulation is defined conventionally as “the use of more than one approach to the investigation of a research question in order to enhance confidence in the ensuing findings” (Bryman 2003) and has too been claimed in Language Program Evaluation theory as the best way to address validity, particularly, to ensure the contextualization of findings thus increasing their credibility and decreasing the potential political interpretation of findings (Norris 2009).

However, focusing the discussion on theoretical insights about challenges and suggesting prescriptions for conducting evaluations has not been sufficient, and the harder evaluators try to make specific choices for evaluation, the harder it is to overcome the challenges that prevent the realization of evaluation’s full potential. As a consequence, currently, a shift of focus is gradually taking place in evaluation theory which highlights the need to conduct research on actual evaluation practice and to metaevaluate, that is, to evaluate evaluations (Scriven 1969).

As theoretical developments in evaluation have advanced, the need to establish better connection with evaluation practice has been called for. It is, in fact, one of the areas that evaluation as a field needs to address in order to consolidate the field (Shaw, Greene and Mark 2006). Metaevaluation can contribute to building a body of empirical evidence to

allow evaluators and anyone involved in evaluation to learn about evaluation practice and to discover what works in what context. Metaevaluation is a concept and a methodology at the same time, and it is the purpose of this section to provide an overview of the development of metaevaluation as a key dimension in Program Evaluation as well as in Language Program Evaluation (LPE).

1.1.1. Metaevaluation: The Latest Advance in Evaluation Theory to Improve Evaluation Practice and Evaluands

As discussed above, evaluation practice is a complex process, and as such, many things may go awry. As Stufflebeam and Shinkfield state, “Due to a host of technical, political, organizational, and psychological complications, many things can and do interfere with and threaten the success of an evaluation task” (2007, p. 646). Moreover, because of previous bad experiences with evaluation, people may feel reluctant to participate in evaluation practices. Accordingly, learning from past evaluations appears as a useful process “to find out how they are actually used and become more sophisticated about and adept at doing useful evaluations” (Patton 2012, p. 131). Proponents of metaevaluation such as Stufflebeam and Shinkfield (2007) advocate the use of metaevaluation as a way to ensure that evaluations provide “sound findings and conclusions; that evaluation practices continue to improve and that institutions administer efficient and effective evaluation systems” (p. 633).

Moreover, since evaluation implies putting the review object under scrutiny, Dahler-Larsen (2006) asks: “Shouldn’t evaluation itself be open to systematic inquiry, just as the policies, programs, and practices that we evaluate are?” (Mark 2008, p. 115). In fact, this question is associated with the argument that “metaevaluation is a demonstration that (at

least some) evaluators practice what they preach” (Scriven 2009, p. iv).

Alkin (2011) stresses the necessity for feedback for evaluators to improve both their skills and their evaluations which can be done via a formal metaevaluation and by resorting to informal sources such as asking questions to either stakeholders regarding the adequacy of the process, the clarity of information reported and so on, and to a colleague ideally with a higher level of expertise, for feedback on more technical matters. Stake and Schwandt (2006), when discussing evaluation quality in program evaluation, claim that

We have no indicator but we have “critical metaevaluation”, the periodic and concluding review of process and product. We can be more confident in the messages of evaluation when such reviews raise challenges to our representations. (p. 416)

In fact, proof of the relevance of metaevaluation is reflected in its inclusion in the most recent and comprehensive compendium on evaluation theory and practice to date, as one of the key elements to be discussed in an evaluation (Mertens and Wilson 2012, p. 51).

But is metaevaluation really an innovative technology within the evaluation profession? Metaevaluation has been proposed by other authors in the past such as Scriven, who first introduced the term metaevaluation in 1969. At a later stage, in Cronbach and associates’ (1980) influential work, when aiming at reforming program evaluation, an important gap was identified: “There is a great need for mechanisms of criticism and appreciation within the evaluation community itself” (p. 359). In their suggested maxims summarizing what program evaluation should strive to do, a critical review of evaluations is represented as a “promising means” to warrant quality evaluation processes. Reviewing evaluation was perceived, thus, as one of the fundamental sources contributing to a renovated picture of evaluation. Since the kind of evaluation reviewing that was being done at the time was infrequent, informal and belated and when systematically approached it

had a “hit-and-miss character”, Cronbach and associates (1980) advocated “penetrating, balanced, extensive inquiry regarding the evaluation enterprise” (p. 359). In particular, they encouraged

[...] further intimate accounts of particular evaluations – their origins, the planning decisions, the operating choices forced by circumstances, and the history of interpretation, counter interpretation, and influence-non-influence. (p. 357).

Furthermore, since the role of evaluators was identified as one of the aspects of evaluation in need of attention at that time, these authors argued that reviewing evaluations meant that knowing in advance that his or her work will be exposed is a motivation for the evaluator “to put his most self-critical effort into each new project” (p. 359). In this view, the evaluator is perceived as a central player who holds a considerable amount of discretionary power over the evaluation process. It is thus implied a deep level of involvement of evaluators in the set of five procedures that form the basis of the formal conduct of an evaluation namely, (1) understanding the evaluation problem; (2) planning the evaluation; (3) data collection; (4) data analysis; and (5) reporting evaluation findings (Nevo 2006).

Since any evaluation exercise implies a quality judgment, which is exposed to internal and external influences, metaevaluation works as an “extradesign procedure [...] that remove evaluators from program advocates or detractors” (Shadish, Cook and Leviton 1991, p. 71). In other words, metaevaluation is conceived as “a strategy to minimizing bias in evaluation” (Fitzpatrick et al 2011, p. 99). Wingate (2010) corroborates this view and further assigns to summative metaevaluation the role of a validation process which adds credibility and “enhances user’s confidence in using the evaluation findings to inform decisions to expand, modify, or cancel programs” (p. 766), and in formative evaluations, increased rigor and reduced bias are warranted by the mere presence of an evaluator who is conducting

the metaevaluation.

As a human discipline, evaluation has a high level of need of managing politics (i.e. power). As Karlsson and Conner (2006) put it, evaluation and politics are “life partners”. In fact, Kushner (2000) perceives programs as fractals of society as a whole since they “expose democracy and its failings” and therefore “each program evaluation is an assessment of the effectiveness of democracy in tackling issues in the distribution of wealth and power and social goods” (p. 32-33). Somewhat relatedly, Cronbach and associates (1980) argue that the main task of evaluation is to help different groups of people in the “policy-shaping community” with diverse perspectives “in the political process to resolve conflicts intelligently” (Cronbach 1982, p. 8). They also argued that while evaluation may be embedded in a context of command, decision-making normally encompasses “a pluralistic community, not [...] a lone decision-maker” (Cronbach 1982, p. 5). Thus, by understanding that various constituencies may be part of evaluation audiences, the evaluator “seeks to survey a range of interests by assuring confidentiality to sources, engaging in negotiation between interest groups, and making evaluation findings widely accessible. The guiding principle is the public’s right to know” (Patton 2002, p. 125). Moreover, MacDonald (1977) identifies three types of political contexts in which evaluation may occur, namely, autocratic, bureaucratic and democratic, and it is interesting to note how Cronbach disregards the bureaucratic category because an evaluator working as a technician with a manager with predefined plans is a scenario in which “programs rarely operate” (Cronbach 1982, p. 35). Therefore, the emphasis in evaluation theory on evaluation’s political nature clearly calls for a robust system of control of evaluation processes in the checks and balances spirit (powers separation) which is a critical aspect dealt with in politics. Metaevaluation would be that equivalent process in the field of evaluation.

Nevertheless, metaevaluation, as in evaluation, is not free from its own challenges and has been the target of criticisms, since not all evaluators agree on the need to foster the development and practice of metaevaluation at this stage.

1.1.2. Metaevaluative process challenges

Several challenges have been identified in any metaevaluative process. The first challenge has been an explicit concern about how little the evaluation community had accomplished in elevating evaluation to the status of theory. As an example, Nilsson and Hogben (1983) argue that more than metaevaluation, what the evaluation profession really needs is a discussion of “meta-theory of evaluation”, which for them implies addressing questions such as why evaluators evade the issue of theory of values by claiming that evaluation should be value-free, and why the task of an evaluator is to provide information, not to make decisions. However, calls to foster empirical research on evaluation practice intended to narrow the gap between evaluation theory and practice have been increasing in the evaluation literature in recent years. In fact, Mark (2008) suggests the use of metaevaluation as a research approach to contribute to increasing today’s scant evidence base of studies on evaluation and in turn to create a stronger link between evaluation theory and practice (Henry and Mark 2003). Furthermore, the study of values has also been addressed in, for instance, the Emergent Realist Evaluation (ERE) model in which a framework of inquiry modes allows evaluators to evaluate values systematically (Mark, Henry and Julnes 1999) and which as explained in Chapter 3 represents the basis of the metaevaluation methodology framework in this thesis.

Therefore, the advancement of the profession is a common and shared impetus and metaevaluation has too important a role to be simply dismissed. Hence, metaevaluation

provides solid arguments for a shift in evaluation, which is being expressed by the voice of experienced evaluation researchers such as Alkin (2004) who believes that

[...] it is time we spent more time doing evaluations, teaching people how to do evaluations, advising on evaluations, reviewing and critiquing evaluations, meta-analyzing evaluations, and in general, advancing the practice of evaluation (p. 166).

To further illustrate this revived interest in metaevaluation, while it was first included in *The Standards for Program Evaluation* (Joint Committee 1974) under “evaluation accountability” (Fitzpatrick et al 2011), in 1994 metaevaluation was formalized as an independent category in *The Standards* (Scriven 2009). (In 1974, representatives from three national professional organizations came together to form the Joint Committee on Standards for Educational Evaluation (JCSEE). By 1981, the JCSEE was supported by 12 North American professional organizations.)

Leading evaluation theorists such as Stufflebeam not only emphasize the imperative need for evaluators to apply metaevaluation to their work and include it as part of every evaluation they conduct (2001a), but also highlight its significance, by developing further its concept and methodologies. As an example, Stufflebeam (2001b) provides the following expanded operational definition of metaevaluation:

The process of delineating, obtaining and applying descriptive information and judgmental information about an evaluation’s utility, feasibility, propriety, and accuracy and its systematic nature, competence, integrity/honesty, respectfulness, and social responsibility to guide the evaluation and publicly report its strengths and weaknesses (p. 183).

Another challenge has been that the practice of metaevaluation has lagged behind due, perhaps, to the fact that the history of evaluation has been one of introspection “with a preoccupation on demarcations and identity” which has distracted evaluators from the study of practice (Saunders 2011, p. 3). As a result, metaevaluation practice has been scarce in the short history of evaluation. However, a testimony of early interest in metaevaluation

is Stake's (1986) published formal metaevaluation under the suggestive title *Quieting Reform* (1986), which is a key study raising very important questions about the actual effect of social science research when applied to human services. Hence, while "informal evaluation happens all the time" (Wingate 2010, p. 771), as an explorative study on the practice of metaevaluation by Cooksy and Caracelly (2009) shows, only a few evaluations are, in fact, metaevaluated systematically. Notwithstanding the slow pace in taking up metaevaluation, an upsurge of interest began to take place in the final years of the 20th century and is becoming increasingly present and relevant in professional discussions which is evidenced also in Cooksy and Caracelly's study showing that most of the metaevaluation efforts seem to have converged between the years 2000 and today, which signals the increased relevance of conducting metaevaluations.

Furthermore, in recent times, evaluation quality assurance emerges as a particular problematic issue in the educational evaluation literature (Cousins and Ryan 2009) due to the influence of the new educational policy trends which may gear evaluation towards purposes other than improvement.

Therefore, in the light of these challenges, besides increasing the research base of empirical studies on evaluation, thus creating a stronger link between theory and practice, metaevaluation studies can become a key contributor to shedding light on evaluation practice: first, to check that evaluation practice is informed by evaluation principles and offers a quality process, and second to raise and help minimize any potential value imbalances.

Taking into consideration that the evaluand under study in this thesis includes a series of evaluation practices conducted at different language programs in Australian universities, the next section of this literature review discusses the tenets of a subset of program

evaluation studies known as Language Program Evaluation (LPE), which also emphasize the use of evaluation and metaevaluation for program improvement and enhancing second language education.

From the discussion in the first part of this literature review, there is little doubt that the improvement of the evaluand and of the evaluation process, particularly in the educational sector, is at the heart of program evaluation and metaevaluation. In the field of languages, Language Program Evaluation (LPE) studies draw attention to the benefits of conducting LPE following theoretically-informed evaluation guidelines to improve language teaching standards and language programs at many levels and even to save the life of programs themselves: thus, the essential consideration “that evaluations be conducted in a principled, systematic and explicit manner” (Alderson 1992, p. 275). Moreover, in this subfield of evaluation, metaevaluation has also been highlighted as an important step to increase research on evaluations of language programs to ensure evaluation is a useful tool for improvement and also disseminate metaevaluation results so that others can learn and take the same initiative, so that metaevaluation becomes an intrinsic part of programs. Metaevaluation in LPE is perceived as an important research path which may help to raise and maintain standards in language education, and represents a stepping-stone in LPE’s theoretical and practical development.

1.1.3. From *Doing* to *Using* Evaluation and Metaevaluation as Instruments for Language Program Improvement

LPE studies have a recent history. They have been mostly associated with English language programs (Watanabe, Norris, and González-Lloret 2009), though other second/foreign language programs have gradually become actively engaged, too. They

began to be treated as part of Applied Linguistics in the last years of the 1980s (Norris and Watanabe 2013) for their great potential to contributing significantly to the development of the discipline as a field of research (Lynch 1996).

Accounts of positive perspectives on the usefulness of PE and from first-hand evaluators' experience which contribute to "promoting effective research and teaching—things we all care about—to people who live on the same planet and don't usually need to be convinced that what we do is important" (Brown 2012, 151), are thought to be beneficial for language educators to become actual *users of evaluation*. Moreover, reflecting on evaluation practices and reporting publicly the aftermath of the evaluation, especially of those evaluations which have yielded positive results, has been considered an activity to be promoted among higher education professionals in languages programs (Norris 2009; Shawer 2013; Birckbichler 2006; Gorsuch 2009).

Consequently, the systematic assessment of the influence of evaluation on the improvement of a program has recently been pointed out as a shift needed in LPE which highlights *using* PE more than *doing* (Norris 2009; Elder 2009; Kiely 2009). Norris (2006, 2009) emphasizes the importance of analysis and reflection when disseminating the results of evaluation in the tertiary languages sector, as this enables language educators to understand all that LPE can offer to the languages profession. In fact, key LPE theorists have claimed metaevaluation in particular as a necessary practice. For example, Rea-Dickins and Germain (1992) defend evaluating evaluations as a feature of any good evaluation process. Alderson's guidelines (1992) to conducting PE include evaluating evaluations as well, which is linked to evaluation utilization, that is, as explained above, ensuring that the evaluation findings are put to use. When metaevaluation is an integrated component of PE and is carried out by the very evaluators who conduct the evaluations, what is achieved is the

improvement of “the evaluation process, and thus contribute to the usefulness and relevance of evaluations.” (p. 299). Richards (2001) also considers reviewing evaluations as an important step of the evaluation process. He defends that this should be done before acting upon evaluation findings to ensure that the evaluation design takes into account the following seven main areas: (1) scope; (2) audience; (3) reliability; (4) objectivity; (5) representativeness; (6) timeliness; and (7) ethical considerations. Kiely and Rea-Dickins (2005), highlight, in Patton’s utilization approach (1997), the need to evaluate evaluations, the implementation of which, in their view, may involuntarily be limited by “the features of practice” (p. 39).

Furthermore, research on LPE practice, in general, and “scholarly traditions of metaevaluation”, in particular, have been called for in the recent LPE literature as a fundamental issue to be considered in order to create an empirically-based body of studies that provide information about actual evaluation practices. Metaevaluation is advocated particularly if PE is to be taken seriously as an integral part of language programs (Norris and Watanabe 2013, p. 6). In sum, the increasing interest in metaevaluation in LPE studies has given it a key role in ensuring evaluation practice is a quality theoretically-informed process. Despite its clear achievements, the relatively short life of LPE makes it stand in a somewhat vulnerable position in terms of its continuation as a subfield within Applied Linguistics. Hence, to truly create a culture of PE within languages programs, claims as to its benefits need to be supported by a number of additional activities, for example the creation of evaluation expertise within languages programs, knowledge dissemination activities, and most importantly conducting research on evaluation within which metaevaluation studies play an important role. Language Program Metaevaluation (LPM) can lead first, to guaranteeing that PE does in fact work for the languages profession (Norris and Watanabe

2013) and second to instilling a proactive predisposition in language professionals to concentrate their efforts on using rather than just doing PE (Norris 2006).

1.1.4. Using Language Program Metaevaluation

Some published reflections on evaluation practice in the context of languages can be found both in higher education as well as in other contexts. In fact, there exist several published individual evaluation experiences of such accounts (e.g. Dassier and Powell (2001), Mathews and Hansen (2004), Rivera and Matsuzawa (2007), Gorsuch (2009) and others). They are indicative, according to Shawer (2013), of a category of research which “marks a move from doing to using evaluation and has been dramatically increasing at the turn of the 21st century” (p. 2886). The focus of the studies by Shawer (2013) and Gorsuch (2009), for example, is to document and offer a reflection about how LPE was used in order to deal with issues as serious as the threat of closing languages programs, as in Shawer’s study, or entire universities as in Gorsuch’s study. In these two cases, the internal members of the program conduct a self-evaluation and provide a reflection revealing what was learnt and calibrate the effects of the evaluation on the program. In fact, Shawer’s study (2012) represents not only a retrospective analysis of an evaluation, but a formal metaevaluation in the field of languages at the tertiary level. This metaevaluation assesses the impact of program evaluation on (a) program improvement and (b) on the professional development of the staff in a language program at King Saud University in Saudi Arabia. Findings reveal that the evaluation did improve both the program elements and faculty and administration’s professional skills. Hence, the positive outcome of this study fosters the idea of the incorporation of metaevaluation as “an integral part of program evaluation to help program stakeholders not only do but also use evaluation” (Shawer 2013, p. 2883).

Another study by Shower and Alkahtani (2012) in the same vein, seeks to align the purpose of their formal metaevaluation with what “many voices have asked” in the field, that is, change the focus from “doing program evaluation to assessing the consequences of the evaluation process for program performance and stakeholders (Byrnes 2008)” (p. 1336). The purpose of the metaevaluation they carry out formally examines the relationship between program evaluation experiences and stakeholder career satisfaction. Researchers compare the influence of *initial* PE experiences on program staff’s perception about PE and career satisfaction with *concluding* PE experiences. Though the correlation found was weak and so was the external validity, the important finding was that “the more program evaluation experiences faculty members have, the less dissatisfaction they show about their career and program evaluation process” (p. 1343). In relation to external validity, the authors state that the study “puts question marks about generalizing this weak relationship into other contexts even similar ones” (Shower and Alkahtani (2012, p. 1343).

Other published evaluation experiences emphasize a wide range of benefits. To mention some examples, PE helped to spot areas in need of attention which otherwise would not have been identified (Morris 2006); evaluation sustained the language requirement in a university campus (Calvin and Rider 2004); evaluation reduced the need to impose external reviews (Mackay 1994) and evaluation contributed to discourage administrators to continue certain cost-saving language instruction methodology when findings reveal a negative influence on student language learning (Sanders 2005). Finally, evaluation helped teachers increase their professional skills, which can raise the professional status of the field and in turn encourage administrators to keep financing the languages (Bernhardt 2006).

Furthermore, since evaluation is a dynamic process, initial impetuses and purposes

may bring about new and unexpected positive academic opportunities. For example, evaluation motivated by institutional accountability requirements, implying a stakeholder involvement, in particular of program staff, in the planning, analysis, reporting and evaluation use of the entire evaluation process. These evaluations may prompt language professionals to inquire about aspects of the program which the original plan did not contemplate, such as student self-efficacy (Gorsuch 2009), or what use students make of the L2 outside of class when promoting life-long learning is one of the principles guiding teachers' instruction (Rivera and Matsuzawa 2007). These evaluation experiences also reveal the wide variety of program layers that an evaluation may focus on, going from an entire foreign language teacher training program (McAlpine and Dhonau 2007) to the comparison of stated and implemented goals (Matthew and Hansen 2004) or even a variety of layers based on different potential scenarios which point out specific educational challenges (Dassier and Powell 2001).

Watanabe, Norris and González-Lloret's study (2009) is particularly instrumental in building an evidence base of empirical LPE studies in the area of higher education. The aim of the research was to learn about the realities of the capability for evaluation in US foreign language university and college programs and to try to find ways to addressing their evaluation needs. Lessons learnt from its findings will be commented in more detail in the Discussion and Conclusions chapters of this thesis since they may illuminate beneficial ways of conducting metaevaluations in the Australian university context, despite the fact that this study is based in the US context.

1.1.5 Fostering an on-going Utilization-Focused LPE evaluation

Evaluation utilization is an area of concern for evaluators in general and so it is for language program evaluators who perceive the use of findings as the ultimate goal of evaluation since utilization is what differentiates research from evaluation (Alderson and Beretta 1992). However, experience proves that the assumption that evaluation's practical aim will be accomplished is not realistic because more often than not certain circumstances prevent findings from being used. While *utilization* is an important evaluation task which was first believed to help at the decision-making stage of an evaluation before acting upon findings, Alderson (1992) cautions that "consideration be given at the planning stage to the dangers of the evaluation report simply gathering dust after it has been submitted" (p. 298). But to factor utilization in the evaluation design requires the negotiation among the parties involved in the evaluation (Alderson and Beretta 1992).

Hence, the instrumental use of findings is considered in LPE essential regardless of the care taken in conducting evaluation activities since "an evaluation that is not used is in some important sense a failure" (Alderson 1992, p. 298). It is because of the significance and weight attributed to the utilization aspect of evaluation that Patton (1978) developed a utilization-focus model of evaluation based on democratic principles which has since been considered the most influential evaluation approach of the 20th century (Norris 2008). In this model, which belongs in the branch of pragmatic evaluation theories, evaluation focuses on the use of findings from the beginning planning stages to the final reporting and ensuring that all stakeholders' perspectives are taken into account. Despite Patton's argument about the second dimension of utilization -- what he names *process use* which, as stated above, consists of the learning that individuals gain through their participation in evaluation (Patton 1978)³, in LPE, the aspect of utilization that is emphasized mostly is the instrumental

use of findings as the following quote implies:

Effective program evaluation in general and language program evaluation in particular can be a powerful improvement strategy when stakeholders and clientele focus on using more than doing evaluation (Norris 2006, p.581)

In the evaluation literature, further essential conditions underpinning this utilization function resonate with the ones Alderson and Beretta (1992) provided for LPEs which emphasize (1) the engagement of intended users in all evaluation processes; (2) the allocation of time and resources to help manage evaluation activities in a practical way; (3) ensuring findings are relevant and credible; and (4) the need to report findings in a timely fashion and communicate them in a meaningful way (Norris 2006).

Therefore, in contrast with the aim of research, “even when use is less direct and immediate, utility of some kind provides the rationale for evaluation” (Weiss 1988a, p. 15). However, it still concerns evaluators to find out that the conceptual (for understanding) and symbolic (for compliance) uses of evaluation are more prevalent than the instrumental (for decision-making and action) use of evaluation (Nevo 2009).

Empirical studies in LPE showcasing good evaluation practices that aimed at utilization show how this might be done. The implications of Watanabe et al’s (2009) study on foreign language departments evaluation needs commented above are considerable. The results of the study led to a workshop entitled *Developing Useful Evaluation Practices in Foreign Language Programs*⁴ aimed at teaching attendees “methods of data gathering, and crucial steps in evaluation projects from forming initial evaluation questions to acting on results” (Ramsay 2009, p. 166) under the utilization-focused evaluation framework. The evaluations conducted by the language educators involved in the workshop in their own programs are thus of special relevance for this thesis because they are all conducted in the higher education setting and illustrate successful evaluation practices.

Table 1 below presents a summary of several key aspects of the evaluations, in particular 1) the impetus that motivated their initiation; 2) the specific purpose of the evaluation; 3) the level of local engagement in the conduct of the evaluation; and 4) the evaluation approach that guided the evaluation. Hence, while most of the evaluations are internally motivated, some had also been prompted or were conducted in parallel for the purposes of accountability or accreditation. To find out evaluations impetuses is relevant for they reflect the variety of concerns and needs that motivate the taking up of an evaluation in a particular program, which in these case studies show not only a wide variety in the nature of those concerns and/or needs but also a variety of the layers of the program on which they focus. And as for purposes, since these are aligned depending on the initial motivation, almost every evaluation presents a different purpose except in the cases at Indiana and at Duke where the purpose in both involved the assessment of student outcomes. Finally, in all of them there was local engagement in every step of the evaluation process including acting on the findings.⁵

7 Case Self-Studies: A Showcase of Intentionally Useful Evaluation Practice in US Higher Education						
College/ University & Language	Internal initiation	Impetus	Purpose	Local Engagement	Utilization Evaluation Approach	Findings Used
Georgetown University, Washington DC German	√	To understand whether unified learning goals were appropriate for and could be successfully implemented in a collegiate curriculum with an academically diverse student body	To gather learners' feedback, past and present, about the appropriateness, level of satisfaction, and perceived learning outcomes at GUGD	All department including teaching staff, chair, graduate students, and faculty	(Program outcomes evaluation) √	√
University of New Mexico Portuguese	√	To update course offerings & attract more students and increase enrolments	To identify whether students had enrolled in the newly created appropriate courses and collect feedback that would help improve course delivery	Program faculty	(Brown's approach 1995) √	√
University of Evansville, Indiana Foreign Languages	√	Since previous internal evaluation claiming the need for improving academic excellence: To reflect and improve the program and disseminate information about program's quality and value	To measure educational effectiveness of department through SLO	Chair encouraged project and faculty conducted	√	√
Duke University, North Carolina Foreign Languages	√	To establish ongoing evaluation and assessment practices within it to validate and improve the foreign language requirement curriculum	To investigate and assess student performance in language proficiency and cultural understanding upon completion of education requirements	Dean + Vice provost + faculty	√	√
Linfield College, Oregon Spanish	√ + Accreditation	To make changes to study abroad program; to continue an ongoing evaluation process to deepen knowledge about study abroad programs and increase capacity to meet institutional accreditation	To investigate causes for unsatisfactory students' proficiency levels upon returning from study abroad	Faculty	√	√
California State U. Monterey Bay (CSUMB) Foreign Languages	+ Accountability	To change Degree Program to ensure students and other stakeholders understand composition of program and are interested in enrolling.	To increase and retain higher number of students as a result of campus-wide enrolment and retention crisis.	Faculty	√	√

Johns Hopkins University, Maryland Foreign Languages	√	To gauge feasibility and utility of a future training program & to address tensions between language professionals vs non-language professionals.	To study the different stakeholder's perceptions of their professional roles and their ideas on professional development for GTAs in FL teaching.	Language Program Directors	√	√
---	---	---	---	----------------------------	---	---

Table 1 Some language program evaluations in Higher Education in the United States

Moreover, the study by Norris et al (2009) of these evaluation accounts shows how in a more educatively oriented evaluation, the very fact of engaging in evaluation leads to changes occurring beyond the planned goals for the evaluation. For instance, professional communication is fostered not only because close collaboration and communication within and between programs ensued “stakeholder enlightenment and deepened program understanding” (p. 217) but also communication flowed vertically between institutional administration and department faculty. An additional benefit from communicating about program understandings may also be the generation of “a common language and a shared conceptual framework to tackle future program issues.” (p. 217). These evaluation projects furthermore increased ‘evaluation competency’ whereby not only evaluators but a range of other audiences involved in the process such as teaching staff and graduate students “were acculturated into evaluation practice” which in turn aids in creating “an evidence-based decision making culture” (p. 216).

The important consideration of evaluation utilization in LPE is reflected also in a number of recently published accounts of evaluations discerning the conditions that either hinder or enhance the usefulness of the evaluation to inform about the effectiveness of a program. The specific recommendations ensuing from the lessons learnt from these internationally-based evaluations are particularly relevant to the context of the evaluation practices studied in this thesis, which significantly fail to meet the conditions that are

proposed. The four most important resemble to a large extent the conditions discussed above which help enhance the utilization dimension of evaluation:

(1) teacher participation and involvement, which helps gain a more comprehensive understanding of what happens in language programs and increases the chance of findings being put to use;

(2) the use of a variety of methods of data collection, which helps capture the measurable outcomes of language teaching on the one hand but also interpret them from a variety of perspectives;

(3) the contextualization of evaluation findings via the use of triangulation (a cross-check of data from multiple sources) of data sources and methods, which helps prevent inaccurate and/or politically motivated interpretations; and

(4) reporting and finding strategic and persistent communication of findings, which helps diverse audiences accurately understand language teaching programs.

A further recommendation is supplied by Norris (2009) who, in studying the evaluation reports, finds that the role of the evaluator plays a very important part in making the evaluations successful. In particular, he points out that the evaluator maximizes the evaluation when

(1) evaluators engage in considerable learning about the language program and its learners, teachers, social circumstances, and educational approach.

(2) evaluators aim at a proactive involvement in evaluation design and implementation of methods that can respond to the intended uses and users;

(3) evaluators construct their efforts within frameworks that make sense to the stakeholders and audiences related to language programs;

(4) they pursue a deep and meaningful relationship with the specific language teaching milieu;

(5) evaluators adopt an educative orientation to their work thus helping language educators to understand the potential contributions to be made by evaluation.

Of course this assumes that the evaluator has the responsibility of determining such design features of the evaluation they are carrying out. This appears not to have been the case in many of the evaluation case studies in this thesis. Instead the design of the evaluation was subject to the constraints imposed by the administrators commissioning the evaluation.

1.1.6. Fostering an Ongoing Evaluation in LPE

Birckbichler (2006), too, appeals to using program evaluation to help those involved in second language programs to respond to external assessment requirements since program evaluation “can offer principled, conceptually sound approaches for identifying the strengths and weaknesses of a program and are a road map for programmatic improvement.” (p. vii). Her approach to evaluation, far from discussing whether outcomes alone or outcomes and processes should be the focus of evaluation, seeks to convey the idea that stakeholders and evaluators should view evaluation as an “ongoing process rather than a summative exercise that stops once the report has been written and distributed. Instead an evaluation should be viewed as a process that assesses the past and current state of a program and that looks towards its future.” (Birckbichler 2006, p. 118).

In the light of this, a whole volume is published offering suggestions on the four broad activities or elements that comprise the conduct of an evaluation, that is, evaluation approaches, data collection and analysis, reporting, and planning for action. In this book, the

recommendations presented acknowledge the variation of contexts in which PE can take place but their proposals are based on evaluation experiences in HE language departments which make the evaluation experience authentic, meaningful and effective.

Thus, in terms of approaches, three proposals are made. First, Costner et al. (2006) make use of the metaphor of 'problem solving' to conceptualize evaluation and propose Polya's (1957) *How to solve it* as a theoretical framework. In the program evaluation literature, the problem solving perspective (Havelock 1971) comes out in Owen's interactive evaluation, an evaluation theory with a strong formative orientation where participants play an important role, particularly those "close to the action" (Owen and Rogers 1999, p. 220), and was conceived as part of a set of innovation and change evaluation models to help organizations deal with problems.

Additionally, Kawamura (2006) suggest an ethnographic approach, in particular, the use of three ethnographic techniques, namely thick description, a balance of emic (insiders' perspectives) and etic (outsiders' perspectives) descriptions of the program, and the adoption of a holistic approach in both data gathering methods and the analysis of the data to avoid misrepresentation of the program. One of the important effects of these ethnographic techniques, in Kawamura's opinion, is that they will "change the perception of program evaluation from a "one-shot" evaluation by outside evaluators to a collaborative and continuous effort by everyone involved in the process." (p. 19). Since evaluation is inherently a social endeavor, it makes sense that communication should have a central role fostering a participatory evaluation model.

Regarding data collection and analysis, the quality of data and appropriateness of analysis are for Dassier, Kawamura and Costner (2006) essential for assessing the effectiveness of a program and act as determining factors on the credibility of the findings

but also on the potential ongoing use of evaluation. They suggest collaborative data collection as well as a wide variety of data types and collection methods such as surveys, proficiency testing, open-ended questionnaires and informal interviews, on the one hand, and qualitative as well as quantitative methods for a well-rounded analysis, on the other. In particular, they describe in detail focus groups and proficiency testing as two useful data collection techniques representing qualitative and quantitative methodologies respectively. While focus groups help gather a diversity of perspectives and opinions, proficiency testing is an essential tool to document the success of a program though they stress the need to cross-examine this data set with other program evaluation data.

The third area presented in this volume is reporting, which according to Lang (2006) must be undertaken in a spirit of openness, and the reporting document should be distributed to a wide audience.

Finally, when planning to act upon recommendations, Birckbichler (2006) highlights the need to pay heed to what Patton (1997) calls the 'personal factor', that is, to identify who is responsible to implement actions for change. Also, when evaluation occurs for the purpose of implementing change⁶ Preskill and Torres' (1999) three steps can be followed which involve (1) identifying and selecting alternatives of action; (2) developing an action plan; and (3) implementing the action plan and monitoring its progress which emphasizes the idea that for evaluation to best serve the needs of a language program it "should be built into the structure of a program, department, or even larger administrative units" (Birckbichler 2006, p. 122).

As an example, Mackay (1994) provides an illustrative case of a system of internal evaluation set up for a group of English language programs using performance indicators as a framework where teachers and directors decide on which components to evaluate but

also the criteria for judging them. For Mackay, evaluation consists of “the systematic collection of information about context, activities, characteristics, and outcomes of individual programs, for use by specific people, to make specific decisions, regarding what programs are doing and who the program is affecting” (p. 142). Therefore, his view of evaluation implies that the information gathered by internal members of a program may be used for the improvement of the program as well as for satisfying the informational needs of people outside the program, such as commissioners. He adds too that establishing an internal evaluation practice in programs would “reduce the perceived need to impose an evaluation from the outside.” (p. 145) whereas the lack thereof “encourages the bureaucracy to impose the only kind of evaluation framework with which it is familiar and comfortable” [...] –evaluations whose focuses, methodologies, questions, and indicators are often alien to the implementation and improvement concerns of the front-line staff of the programmes concerned.” (p. 145).

1.1.7. Nevo’s *Dialog Evaluation Approach* (Nevo 2009)

Finally, a set of recommendations about evaluation design and the implementation of evaluation findings that are relevant to the context of this study are contained in the Dialog Evaluation Approach (Nevo 2009). They include:

1. Recognition that while not everybody will have the same authority they will all be equal in that all will have something to learn and something to teach
2. Engagement by all in systematic data collection which implies on the one hand that the internal team will need to be acquainted with qualitative and quantitative methodologies, and on the other that focus on relevant issues and pertinent data is essential since choosing trivial issues will hinder dialogue

3. Parties of internal teams and external evaluators must cultivate respect and trust to establish good communication channels for the flow of information to take place;
4. Fairness can be accomplished by attending to legal and ethical issues and also making clear to internal and external evaluators about evaluation purpose, benefits, costs and who is responsible for paying those costs
5. Responsibility for handling the consequences of the evaluation which should be shared between internal and external teams. In Nevo's view, a program or a school left alone to handle recommendations will most likely lack the necessary motivation to "move into a serious dialogue for [...] improvement" (p. 301).

1.1.8 From Evaluation to Metaevaluation: Summary

To conclude this section, from the instances of formal research studies on evaluation published so far, metaevaluation in LPE seems to be still in its early research stages.

However, it is also clear that metaevaluation is gradually attracting the attention of LPE researchers. The experience of the advantages that metaevaluation can bring for the improvement of both the evaluation practices themselves and for the language programs, including program stakeholders, signals a trend in LPE that fosters *using* rather than *doing* evaluation in various educational contexts including higher education.

Additionally, an increase in bureaucratic types of evaluations has drawn the attention of LPE theorists who see the bureaucratic approach to program review as the one most common where programme personnel have not been invited to contribute to the design or focus of the evaluation study. That is, in situations where the evaluation questions have been formulated virtually entirely by the funder or the bureaucracy in isolation [...]. (Mackay 1994, p. 143)

which fails to contribute to a perception of evaluation that fosters learning about policy functions and professional practices (Cousins and Ryan 2009) but rather to a culture of quality management and compliance. For Kiely and Rea-Dickins (2005), compliance with quality mandates may potentially trigger a certain level of stimulation of institutional evaluations or even institutional self-evaluations, but it can also present two drawbacks. The first one involves the issue of relying on static (documentation) rather than on dynamic (what actually happens) characteristics of programs. The second one questions who decides what is in an evaluation and highlights the risk of management determining curriculum since this detracts from teachers' confidence in the program they work in and may have a negative washback effect on curriculum and the learning experience. In Britain, a study by Norris (1998) focused on curriculum evaluation identifies six approaches used, four of which "represent forms of monitoring to ensure compliance with mandates": performance indicators, self-study, expert or peer review and inspection. It is interesting to note that even self-reviews, in quality management systems, have become externally regulated processes "with a significant bureaucratic dimension" (Norris 1998, pp. 209-14). Hence, the imperative need to consider issues of power distribution and to identify the locus of control since in bureaucratic evaluations particularly "the power base in many evaluation cases resides with the evaluation sponsoring agency or the evaluator/evaluation team" (Kiely and Rea-Dickins 2005, 216). Moreover, Mackay (1994) comments on the "dissatisfied, frustrated, confused, and even hostile" mental state that bureaucratic evaluations leave program staff in and wonders about the effectiveness of such evaluations to undertake program improvement (p. 145).

Therefore, the political dimension of evaluation is brought up also in LPE theory, thus raising the additional metaevaluation function of checks and balances as a critical one besides the research and quality improvement dimension in the practice of LPE as it is in PE.

But what is the state of the art in the Australian context? Are LPE and LPM being developed? To what extent? The answers to these questions and related ones can be found by paying attention to one of the fundamental issues in evaluation: context, which is the purpose of the concluding part of this first chapter.

1.2. Context and Metaevaluation

This section explores the context of evaluation practice in the Tertiary sector of Australian universities. The analysis of background information will stress the need to foster evaluation-informed practices, including metaevaluation, as highlighted in the previous sections of this chapter. Three contextual dimensions are examined in the next section from which further research inform findings in this thesis more thoroughly, In line with metaevaluation theory:

(1) Context 1: Language Program Evaluation and Metaevaluation in Australia, in order to find out the kind of work done so far and future directions in terms of their level of use as an integrated component to helping language education improve in the Australian context.

(2) Context 2: Evaluation developments in Australia, in order to set the scene for understanding the role that has been assigned to evaluation and the challenges that evaluators have faced in conducting evaluation in this national context.

(3) Context 3: The reforms that the Australian Tertiary System has undergone since the end of the 1980s, hand in hand with the implantation of the New Public Management internal policy and organizational framework in universities, which have come to shape the tertiary sector as a business model (Marginson 2009).

At the same time research from within these contextual dimensions will inform the findings that result from the research-oriented metaevaluation in this thesis. Thus, the study of context is divided into three main parts:

Context of Metaevaluation 1: Language Program Evaluation and Metaevaluation in Australia

Context of Metaevaluation 2: Historical Developments in Evaluation in Australia

Context of Metaevaluation 3: Understanding the University New Public Management

Business Model Built Upon Neoliberal Ideology

As seen throughout the previous literature review discussion, evaluation practice is believed to be conditioned by a number of factors, including context. This argument has been suggested by evaluation theorists who have emphasized examining contextual factors extrinsic to the evaluation exercise itself, since an understanding of them is essential to increasing honesty and balance in program evaluation (Cronbach and associates 1980). In the same vein, others highlight the relevance of context (Henry, Mark and Julnes 2000), particularly in evaluation theory that focuses on higher education (Grayson 2012, Premfors 1989); or specific evaluation methodologies which encourage researching context prior to conducting evaluation (Alkin and Christie 2013; Darling-Hammond, Wise and Pease 1986). In fact, in metaevaluation theory, context, too, is discussed as an important aspect to examine prior to the metaevaluation (Stufflebeam and Coryn 2014).

During the process of this research study, the category of context has emerged as one of particular relevance, in parallel with recent metaevaluation studies conducted in different international higher education institutions which show how the policy context ruling universities today has come to characterize evaluation in distinctive ways (Saunders, Trowler and Bamber 2011).

Since the purpose of this doctoral metaevaluation is to explore and characterize evaluation practice as it happens on the ground, beyond the consideration of the traditional evaluation fundamental issues (e.g. purposes, methodologies, evaluation approach, etc. detailed in Appendix 4) it seeks to gain a clear picture of the context to shed light on findings. In particular, it researches various aspects of the context associated with several institutions in the Australian tertiary system from where data were obtained.

1.2.1. Metaevaluating Context 1: Language Program Evaluation and Metaevaluation in Australia

PE is a strong research area within Australian university Departments of Education, including existing Evaluation Centres (e.g. The University of Melbourne), being an active area of research⁷ although there are scant studies on evaluation in the Australian higher education sector.

Regarding LPE, despite the lack of an established practice of evaluation in university language departments to date (Martínez Marco 2013), and the fact that research on LPE has not been a priority at a university level, a few instances of relevant LPE research in other Australian educational contexts can be found. For example, McNamara and Eisikovits evaluated a late bilingual immersion program in a Jewish high school in Melbourne between 1990 and 1991. Hill and McNamara (2003) reported on how measurement of student learning outcomes is used to evaluate curriculum initiatives and policy in primary schools. A more recent example is the published evaluation study by Elder in 2009, which reports on and provides a retrospective insight into an evaluation of a series of state-funded bilingual programs in primary and secondary schools in the state of Victoria.

Australian Language policy is another area of language education which has received extensive attention recently. It provides key background information to understanding the Australian context in which evaluation is taking place at the present time. In this sense, studies from Lo Bianco (1994), Lo Bianco and Slaughter (2009) are revealing. Additionally, Martín (2005) provides a panoramic account of the current situation of languages programs in Australian universities in relation to the historical developments of its education system in relation to second language teaching.

A coherent language policy would help mitigate the ongoing crisis that language departments in tertiary institutions in Australia, as acknowledged in the National Languages Summit of 2007. The critical situation in Australia for language education was recognized at the summit as a result of a gap identified between “the rhetoric and reality; between intention and effect”, thus calling particularly on “federal, state and territory governments to develop an agreed national languages capability target for a significant majority of Australians to attain second-language proficiency by 2020.” (Communiqué, National Languages Summit, National Press Club Canberra, Thursday 7 June 2007). In fact, there are signs pointing towards a re-stimulation of evaluation practice as seen in the principles for the Languages and Cultures Network for Australian Universities (LCNAU). This institution acknowledges the importance of a) on-going evaluation practices for university language programs; and b) the reflective and formative aim of these evaluations, which resonates with the argument to pursue LPEs that are educational in nature. Furthermore, Australian applied linguists such as Scarino and Liddicoat (2009) stress the important role of evaluation as (1) “an ongoing process of building understanding of professional work” (pg. 87); (2) a process that helps reflect the stance of the teacher; and (3), “an integral part of the process of curriculum renewal” (p. 87).

Metaevaluation, both in PE and in LPE, has been called for, as we shall see, to address an important gap in evaluation theory, but also to attend to the practical inherent aspect of evaluation which is its usefulness for the improvement of the evaluation process and for the betterment of a language program.

There have been evaluation trends and specific challenges influencing Australian university life in general and evaluation practice in particular, which have led evaluation to have its idiosyncrasy (Sharp 2003). The relevance to research on these trends and challenges in this metaevaluation lies in that they have triggered the development of a 'narrative'⁸ derived from managerial principles. This narrative emerges in the data analysed in this thesis, either on the surface or at a latent level. The application of its principles has created a number of serious risks in Australian universities. The discussion below explains the origins of such trends and their effects on evaluation practice after examining key historical evaluation developments in Australia (Context 2) and the characteristics of the Dawkins Reforms underlying a neoliberal ideology for the Australian Higher Education sector, exploring their effects on evaluation processes within universities (Context 3).

1.2.2. Metaevaluating Context 2: Historical Developments in Evaluation in Australia

Evaluation in Australia has been traditionally associated with government programs which inevitably subjects it to governmental ideologies. While drawing on international evaluation traditions such as the ones of the US and Europe, evaluation in Australia reflects a mix of both outer influences but also in-country developments which have shaped its distinctive character. According to Owen (2003a), a renowned Australian evaluator, Australian managerial evaluation origins differ from the academic origins of evaluation in the US. Historically, evaluation practice in Australia reveals a trend towards the

professionalization of evaluation which is analysed in the next section under *The Process of Professionalization of Evaluation in Australia*. This might have occurred as a result of both the institutionalization of evaluation by governments and under the influence of international evaluation trends. The methodological tension between accountability forms and more open participatory approaches leads the discussion into the description of the second trend in evaluation in Australia, that is, Evaluation for Accountability versus Evaluation for Improvement in a Managerial Culture, since evaluation has been embedded in a managerial framework which has been informing Australian public institutions since the end of the 1970s.

1.2.2.1 Process of Professionalization of Evaluation in Australia

Attempts to institutionalize and formalize evaluation in Australia have occurred mainly through the following three reports as a result of decisive parliamentary inquiries to implant evaluation as a bureaucratic tool (Sharp 2003):

(a) The Coombs report (1976) identified organizational diagnosis and benchmarking as two essential aspects for improving public administration.

(b) The Baume report (SSCSW 1979) remarked upon the absence of a culture of evaluation in Australia and provided prerequisites for and a definition of program evaluation. In particular, it emphasized the accountability of programs for resource allocation. Interestingly, existing evaluation work on agricultural extension studies and the evaluation impetus to institutionalize evaluation in some education areas was not acknowledged in this report. The case of the evaluation of agricultural studies is significant because critics of work in that field have argued that the focus on methods in them is associated with a lack of attention to context, self-evaluation or critiquing evaluations. If the

improvement function was not activated as a result those evaluations there may be a potential relationship with lack of a role that metaevaluation may have played in redirecting the evaluation focus.

(c) The Reid report (Review of Commonwealth Administration 1983) came out in an era where managerialism was in the rise, following the same trend in other Organisation for Economic Cooperation and Development (OECD) countries such as Canada, the United Kingdom and New Zealand. The urge in this report to implant strategies such as management by objectives and performance control led to the creation in 1984 of the Financial Management Improvement Project (FMIP) and the 'Evaluation Strategy' to systematically evaluate all Commonwealth-funded programs. As Cuthbert (2011, p.133) argued, 'Managerialism, as an internal organizational framework in universities, is a set of constituted and constitutive social practices involving actors inside and outside the institution, especially senior managers. It is associated with new public management, emphasizing specification of outputs and targets, performance measurement as a means to its management, and business inspired practices such as contracts for service, increased competition between HE providers, and a quasi-market framing of students as customers.' Moreover, NPM (New Public Management) was first established in New Zealand in the late 1980s and was rapidly extended to other countries, particularly in the Anglophone sphere, among which some similarities were identified, Australia being one of them: "Early parallels were drawn among them, and Australia, New Zealand and United Kingdom were grouped because they adhered more to the precepts of NPM than other OECD countries (Hood 1996a). At the peak of the OECD's fixation on NPM, Anglophone experiments were upheld as the ideal." (Halligan 2011, p. 84).

All three of the reports we have considered here confirm that the institutionalization of evaluation as a bureaucratic tool meant an economic check to control how public money was spent and allocated, that is, for economic efficiency. In parallel to these developments, practitioners of evaluation gradually forged an evaluation profession with important contributions such as the creation of The Australasian Evaluation Society (AES) in 1986. According to Sharp (2003) in 1982 when the first National Evaluation Conference (NEC) took place with the purpose of assessing in retrospect the state of the art, several pointers were suggested about evaluation developments and practice. One of them was the need to create pathways of communication for those practicing evaluation and more importantly that evaluation trends indicated a resonance with economic national strategies⁹.

Sporadic recurrences of self-evaluation and the recognized need and posterior development of standards for good practice contributed to the idea of professionalization advance. In fact, the main standards of the AES, namely, Guidelines for Ethical Conduct of Evaluations (1997) followed by a Code of Ethics (2000), were an essential contribution for a new discipline such as evaluation as well as an illustrative example of “a formative way to generate a feeling of belonging to and identity with a specific professional community” (Bustelo 2006, p. 441).

Evaluation, thus, seems well on the way to becoming a well-established profession in Australia. With a useful and an important purpose, accountability evaluation is to ensure that resources are well allocated and spent in programs funded with public money. This government principle reflects a limited view of evaluation as a mere efficiency tool. In the realm of universities, the area under study in this thesis, a question remains as to the extent to which Australian professional evaluators are involved in evaluation processes and/or whether evaluation practice in universities is guided by evaluation guidelines as promoted

by the AES, given the scant literature on evaluation in HE in general and in Australia in particular.

1.2.2.2. Evaluation for Accountability versus Evaluation for Improvement in a Managerial Culture

The dichotomy between evaluation for accountability or for improvement purposes became a lively critical debate during the 1980s, when a corporate management ideology was installed in the government and other public institutions. In corporate managerialism, accountability is a key concept which entails the use of system wide indicators to secure that the system or programs in which these operate work efficiently and effectively.

While Australian evaluation practitioners and researchers such as Owen (2003a) recognize that accountability is an inherent characteristic of a democracy committed to openness and access to information, the emergence of a new age of evidence-based decision making (Owen 1991, 2003b) has contributed to promoting an image of evaluation as an instrument for accountability. This orientation, according to McTaggart et al (1991) comes from internal influences such as training in psychometrics as well as “legalistic blue ribbon panels of enquiry” (p. 124) reflecting Australia’s institutional British heritage. Also, the active commitment to *internationalism* is counted among external influences, which made the Australian government follow closely US moves either directly or indirectly. For example, Reagan’s governmental policies such as budgetary constraints directly inspired “hierarchical efficiency-focused evaluation in the name of ‘scientific’ program management in Australia” (p. 129); or indirectly through international associations such as UNESCO or OECD where issues about the US became issues of relevance in Australia. One particular

feature of evaluation of government programs under the auspices of globalization is that it tends to use criteria based on economic terms.

In the past, accountability evaluation was contested by fostering the use of participatory approaches, even though, in accountability evaluation, 'participation' is traditionally used as a co-optive technique not warranting the neutrality and objectivity needed for accountability. The debate about approaches in the 1980s in Australia reflected the concern about how certain approaches suited certain kinds of institutions and organizations. As an example, participatory approaches are not just difficult to establish in hierarchical institutional contexts, such as NPM, as they may threaten the very hierarchies which frustrate their efforts. These approaches fell in disfavor and were made problematic for various reasons. As an example, participatory approaches challenge central concepts of corporate management, particularly, (1) the privatism of bureaucratized systems which challenges the process of making information public; (2) the traditional view that "evaluation practice is the domain of experts who work in central bureaucracy, or in universities and institutions of research" (p. 128); (3) the complaint presented by administrators and program practitioners regarding the workload involved.

In sum, McTaggart et al's research on evaluation developments in Australia in the 1990s leads them to conclude that the contestation between participatory approaches and corporate management evaluation remained unresolved. This was due, for instance, to the variation of evaluation practice contingent upon different evaluation needs and the methodology becoming problematic as a result of the notable politicization of evaluation practice. It was obvious as well that the motivation of Australian governments to apply US economic policies, which were contemporary to and of the same nature as Thatcher's economic measures in Britain, continued as vividly in the 1990s as they were in the 1980s.

The attempt from Australian governments to institutionalize evaluation in public administration within the managerial framework succeeded to the extent that it was mainly associated with accountability, thus dramatically reducing the scope and meaning of evaluation. At the same time, the grassroots development of evaluation as a professional activity in Australia brought to the debate the importance of using evaluation processes for improvement challenging a trend to understand accountability evaluation as an economic control instrument alone.

The spread of NPM principles has changed how the Australian Tertiary Sector operates. In new managerialist ideology, according to Trowler (2011), public universities are required to

- (1) find their own sources of funding;
- (2) have a business and commerce orientation in their mission;
- (3) reconceptualize education as a private rather than a public good;
- (4) push participation rates while reducing resources for teaching;
- (5) establish monitoring mechanisms, which have generated an “audit culture” in line with “the government machinery for an ‘evaluative state’;

(6) have more openness and accountability to increase competitiveness and raise standards. This means the publication of data about performance which has led to the use of external measures such as league tables and international rankings to measure universities’ achievements.

Universities’ priorities have changed accordingly, arising from implementation pressures coming from the government associated with neoliberal economic policies, as well as from the managerial hierarchy within universities themselves. In the top-down performance management context of NPM, evaluation plays the role of a control

instrument to ensure NPM principles and priorities are put in place. The alignment of evaluation purposes with accountability of programs, departments or schools under the NPM policy framework brings up a number of risks discussed in context 3 below, which become particularly relevant due to

- a) a focus on (mainly financial) accountability, which undermines the improvement function of evaluation
- b) evaluation processes are void of evaluation principles and standards, which puts into question the quality of the evaluation process itself, thus, the validity of both the process and results
- c) the lack of accountability to stakeholders on the part of those managing reviews opens the door to an increased politicization of decision making derived from evaluation.

1.2.3. Metaevaluating Context 3: Understanding the University New Public

Management Business Model Built Upon Neoliberal Ideology

While one could accept that the tertiary sector may have been in need of change, improving the sector did not require its destruction. (We are assuming here that improvement is the objective of reform). (Meyers 2012, p. 10)

In this section, a description is presented of the most notable features and implications of the reforms and university model introduced in Australian universities, which are relevant for understanding the context under which reviews are being conducted in that setting. More precisely, Section 1.2.3.1 Embracing New Public Management Principles to Overcome a “Malaise” looks at the reasons why the reforms were adopted to tackle various problems which had been identified in universities for some time and which called for

important changes. Section 1.2.3.2, The Impact of the NPM Policy Context on Educational Quality, discusses the effect of the adoption of the NPM policy context and the particular style and character adopted in Australian universities on educational quality. Finally, Section 1.2.3.3, Strategic Policy Use of Quality Assurance Evaluation in Australian HE Preventing Improvement reviews the different approaches preventing quality evaluation implementation in Australian HE.

1.2.3.1. Embracing NPM Principles to Overcome a “Malaise”

Public universities in Australia have been undergoing restructuring and reforms as a result of ongoing neoliberal governmental policies which started in the late 1980s by Labor Party Education Minister John Dawkins (Harman 2006). Published discussion voicing the effects of the Dawkins reforms and the incorporation of New Public Management in Australian university life are referred by quite a few Australian authors (e.g. Coaldrake and Stedman (1998), and others) identifying both positives and negatives of these reforms. As an example, Meyers’s (2012), *Australian Universities: A Portrait of Decline* depicts a vivid picture of how these reforms have created fragmentation in the sector. Also commented on in a study conducted at the Centre for the Study of Higher Education (CSHE) at the University of Melbourne, this fragmentation could be attributed to, at least in part, the risk of misrepresenting management. The strong adherence of university senior managers to new managerialism may run counter to the intrinsic goal of supporting those who are managed: “If they overuse rationalistic analysis, targets and key performance indicators as ‘weapons’ to respond to the ‘attacks’ they face, they may reinforce the very problem which causes the pressure” (Cuthbert 2011, p. 138). Meyers’s book portrays an account of the impact of the reforms since the Dawkins era at the end of the 1980s and continued by

subsequent governments, which has resulted in continuous failed attempts at finding interested publishers. Another significant work with a less blunt commentary than that of Meyers, contains a number of articles written by different Australian academics and edited by Coady (2000)¹⁰ with a diversity of views offering what the contributors consider a necessary 'conversation about values, means and directions' of the current Australian university model.

Among the reasons for the success of the reforms, the support by various groups, namely the public, the higher education community, the press and the business community are pointed out (Harman 1989). They clearly indicate an overall consensus about the need for important changes in the HE sector: universities' financial independence from government requiring funding from private sources; turning HE into an export industry through the recruitment of overseas full-fee paying students; and universal access of universities to all have been perceived as positive outcomes (Murray and Dollery 2006). Notice how the characterization of these changes as positive, denotes a readily open mentality to embracing key NPM principles.

According to Marginson (2009), new funding arrangements and the implantation of a business-oriented management system, emulating the New Public Management British and New Zealand models, has since the Dawkins reforms era operated within Australian universities (Harman 2006). Its key purpose has been to tackle what had been perceived as "a deep seated malaise in higher education" (Harman 1989, p. 28). This significant change has come to characterize the culture of institutions as "businesslike, managed, transparent and performance focused" (Marginson 2009, p. 6). The two main traits characterizing NPM, according to Ryan and Feller (2009), imply more flexibility and improvement, including the following: (1) the decentralization of decision making from governments would allow

universities to flexibly adapt to changing circumstances and respond in more efficient and effective ways to those circumstances; and (2) the encouragement of competition in the public sector would increase the incentive of universities to attain high levels of performance for best practice.

The favourable implementation of the current model is what Harman (2006) calls positive outputs produced by academics during the transitioning years of Australian universities to becoming corporatized. The findings from existing surveys from the early 2000s with responses from science and technology academics on the one hand, and social scientists on the other from two different Australian university groups are revealing. Harman found that academics have shown an impressive degree of adaptability to the new environment; a high degree of involvement in activities related to issues they initially had shown an opposition to, such as recruiting fee-paying students; a high commitment to teaching and research, the quality of which participants claim not to have declined; a strong work ethic; and an active participation in collaborative projects to attract external research funds.

However, the interpretation of some of Harman's findings reflects a certain degree of naivety. As an example, when participants are asked to value their own work in terms of educational quality or about participants' level of interest on core academic activities, it is illogical to expect an objective response from a participant who is judge and interested party at the same time. Moreover, Australian universities are evaluated yearly against a performance-based system which measures their outputs as part of the formula for government to determine allocation of institutional block awards for research (Ryan and Feller 2009). Academics, thus, are pressed to conduct research and publish for universities' financial relief, which assumes particular relevance at the time of the academic

Performance Development Framework (PDF) review. Under this policy context, it seems irrelevant to ask academics about whether their professional activities are or are not driven by an intrinsic interest when their jobs depend upon an annual review informing their line manager's decision.

The PDF review is an ongoing annual process whereby academics are evaluated against research performance, and other indicators, since "tenure of academic staff no longer exists in the Australian tertiary system" (Meyers 2012, p. 116). In fact, in Australian universities a "pressuring atmosphere" is leading to

ongoing scrutiny of scholars' permanent contracts often under changing rules and regulations as well. In fact, annual reviews are designed to set up agreed objectives and aims with the expectation to attract funds, not always strategically aligned with the advancement of a specific discipline (Martínez Marco and Pérez de León 2019, p. 98).

Additionally, Harman's study doesn't provide an in-depth view of other potentially damaging consequences impacting Australian universities today which can, in turn, affect the future of the whole tertiary system. On the one hand, the current working conditions of academics, as shown by a study at the Centre for the Study of Higher Education (CSHE) at the University of Melbourne (Bexley 2013) shows a decline of academic work caused by ad hoc policy leading to new NPM-related work roles. Due to various reasons, such as an overload of administrative tasks for quality assurance on faculty, the quality of academic work has been considered a failure in the new model (Murray and Dollery 2006). Moreover, in the current performance measurement system of both individuals and institutions, a key concept in NPM framing Australian universities, besides unveiling validity issues, particularly the validity of criteria and the use of specific measures, the subversion of the goal of improving quality is apparent. In Ryan and Feller's (2009) words, this leads to "the situation where more measurement of quality may lead to everything else but quality" (p. 183).

Butler (2004) adds in respect of research publication quality that “journal quality, visibility or impact is not taken into consideration when funding is allocated. Therefore, there is little incentive to strive for publication in prestigious journals. The rewards are identical” (p. 394). Bleiklie et al (2011) consider that the concepts of accountability and transparency under the NPM framework are contributors to the new role of academics in Australian universities and their associated quality assurance processes:

The existing lack of stability of academic contracts, both short and long term, added to the frequent changes in academic planning, forced to be adapted to both new top-down strategies from the institution and/or from funding or government bodies, have transformed the traditional way colleagues conceive their academic roles. In this context, it may be argued that academics have been turned into “subordinate workers under constant pressure to produce and bring in fresh funding: They [academics] are expected to and do spend more time on funding acquisition, writing research proposals according to specified formulas including work packages, deliverables and deadlines. They also spend more time reporting on their activities as part of internal reporting, quality assurance and budgeting procedures at their own institutions where the activities and productivity of every individual academic now affect the funding available for their own research group, or their own department or unit within it. These reporting procedures are making the contributions of academic units, but also of individual academics publicly available and visible.” (Bleiklie et al 2011, p. 21)

A major concern as a consequence of the insecurity and conditions of employment lies in the fact that they are giving substantive arguments for people to consider either to move out of academia or to find a career at an overseas university (see Bexley 2013). The pressure for universities to attract funding has therefore reconceptualized the role of academics in significant ways. The main concern of academics is increasingly to ensure their work is aligned with universities’ priorities, focusing primarily on funding. Therefore, it seems that the ‘malaise’ that the reforms were seeking to remedy, has given way to another “malaise that is becoming ubiquitous across the academic workforce” and which is not being mitigated by the “intrinsic appeal of the scholarly activities” (Bexley 2013, p. 103).

The emphasis on accountability and transparency processes, within which reviews would be located, has resulted in a significant reallocation of resources. The reporting activities which universities put together for the purpose of accountability are creating the justification for not only misallocating resources but also “for it to be ‘self-serving’ for those in managerial positions” (Murray and Dollery 2006, p. 486), since there are no mechanisms in place for managers to account for how they go about managing those resources. Thus, it could be claimed that accountability is increasingly exercised only from the bottom up. In this regard, Martínez Marco and Pérez de León (2019) highlight how economic efficiency has become the main principle not only permeating academic work, but being used as the justification for the lack internal mechanisms of power control of those in managerial positions, and as a principle which is superposed over other priorities as important as the use of evaluation for the purpose of academic improvement:

In this well-established NPM culture, it seems that, idealistically, the actual checks and balances of all processes are built in within the actual process itself: an incessant search for efficiency proves the intention to make a good use of public funds and so this principle should be prioritized in all academic processes. While all good administrators are also very efficient, it does not imply that in order to achieve long-term efficiency, reasonable for the quality of teaching and research, short-term efficiency should not be questioned. Actual evaluations and other external mechanisms of self-reflection and academic improvement do not seem to be essential priorities. (102)

Governance is another area which has been considered negatively affected by NPM policies. The decision-making processes, the result of market reforms and adoption of NPM, translates into low academic self-governance in formal and institutionalized decision-making. This sees the university’s administration as the exclusive decision-making authority with the Academic Senate having a mere advisory role and only with power if delegated by the university Council. However, some researchers seem to be unaware of this feature as intrinsic to NPM. For example, Coady (2000) find abusive the “authoritarian, top-down,

cursorily discussed” decision-making processes which “tend to be hostile to diversity and variety” (p. 21), though this perception may also have been given by the characteristic decision-making styles adopted particularly by the management hierarchy in Australian universities. In fact, low academic self-governance will endure as long as universities are required to meet national productivity and economic goals (Traveller 2014). As an illustration, Meyers (2012) stresses not only how the opinion of management is superposed on “academic judgment on educational matters”, but even how academics have no educational authority to influence the determination of resource and organizational priorities. Therefore, in this setting, it would be logical to expect a top-down approach to reviews in universities.

The ‘technical-managerial’ character of accountability in contemporary higher education systems, inextricably inseparable from evaluation, has also raised the issue of the distinctive role of evaluators for this context (Ryan and Feller 2009). In fact, Bleiklie et al (2011) have characterized evaluators’ role in the NPM university context as that of “academic gatekeepers” coming out of internal and cross-institutional elite networks created as a result of the weakening of academic self-governance:

academics sitting on review and selection committees, reviewing papers, selecting projects, and making authoritative judgements on the quality of institutions or disciplines. The impact of their decisions will increase and they are likely to constitute a new academic elite. [...] The position of the members of this new elite is based not just on full professorship and similar academic top- positions, but on network positions gained through participation on academic peer review panels of all sorts, research funding panels, evaluation bodies, hiring committees, editorial boards and so on [...]. Individual members would usually acquire the positions that make them elite members based on research reputation. Within individual universities such elites may be highly influential, at the same time as rank and file academics find themselves in a politically gradually weaker position [...]. NPM policies have contributed to formalising new kinds of policy networks related to external research funding mechanisms, evaluation and accreditation agencies, and institutional governance. (Bleiklie et al 2011, p. 24, p. 26)

Additionally, a validity question in connection to review committees has been posited when their expected effort in collecting and examining data is overridden by “opinions based on their preconceptions and the need to achieve a particular objective” (Meyers 2012, p.140). According to Meyers’s claim, evaluations operated under management in Australian universities runs the risk of invalidity since evaluators may be selected according to their ready predisposition to steer a review in the direction of the specific interests of managers, who without any checks and balances, use the review as an instrument to align a particular program, school or department with specific priorities.

After more than two decades, the consequences of top-down reforms and NPM are beginning to become apparent. However, instead of deterring it, the reading of these consequences seems to have led to yet more NPM regulation.

The Bradley Review of Australian Higher Education (2008) announced a funding approach based on enrolment numbers and performance (King and James 2013) which exerts increased accountability pressures on universities. Along with new funding arrangements, the Bradley review has made universities enforce a series of new measures such as widening higher education participation; improving equity and removing restrictions on the number of student places, which manifests a move towards more deregulation. With funding to universities gradually falling to a 40% by 2005 (Harman 2006) institutions had to strategically find other sources of income, while attempting to cope with university core mission tasks such as teaching and research. As of late, however, intense financial cuts have continued to pressure universities to the extent that “the public funding of higher education as a proportion of the Gross Domestic Product (GDP) has trended downwards to the point where it is now only two thirds of the OECD average” (Marginson 2013, p. 7).

Related risks for academic quality include the impact of NPM in research and teaching quality only measured by specific limited data such as student evaluations.

1.2.3.2. The Impact of the NPM Policy Context on Educational Quality

Harman's finding showing respondents' view about the quality of teaching and research not being affected by the new system contrasts with other more recent academic sources which have identified challenged educational quality as a key current concern. In the CSHE study, the risk of the decline of educational quality in universities is pointed out, associated in part with findings showing that casualization has increased to 68% (Bexley 2013). Bexley (2013) also points out that more than half of individuals who are hired on a casual basis are over 40 and do not study, which demystifies the assumption that casual positions are filled by Research Higher Degree (RHD) students. Furthermore, the new focus for universities provided by the Dawkins reforms were driven by concerns about academic standards, demands for greater accountability and the alignment of university studies with professions, among others. As a consequence it has conferred on the quality assurance system a new orientation: being a policy instrument aimed at ensuring that "institutions and systems have in place mechanisms for review and assessment and for renewal and improvement [...]" adding "much more emphasis on external scrutiny, on seeking the views of employers and graduates and [...] making the results of assessments more widely available (Harman 1998, p. 333). A lack of an established evaluative cycle in the Australian tertiary sector as far as languages education is concerned can be identified (Martínez Marco, 2013), but it is also clear that evaluation in Australian universities is for the most part tied to Quality Assurance (QA) processes mandated by governmental and/ or institutional regulation (Krause 2011). However, while a tradition of peer review and

external evaluation in Australian universities is known (Thomson-Whiteside 2013; Massaro 2013), research on evaluation practice in HE in Australia reflects to a great extent a sector “replete with many forms of evaluation but which has yet to become a focus for research into evaluative practices” (Saunders 2011, p. 1).

Furthermore, Quality Assurance requirements from which most reviews stem count among the NPM features adopted by various Australian systems and institutions and represent a response to economic moves such as “downward pressures and efficiency drives in public spending, concerns about international competitiveness, and measures to enhance skilled migration” (Marginson 2010, p. 5)¹¹. Next to Quality Assurance, Marginson (2010) also mentions other adopted NPM features pointed out above such as corporatized institutional governance, performance measures of staff, and professionalized and strategically focused executive management (p. 5-6). In fact, in the early 1990s, the tendency was to use quality assurance processes in Australian universities “as a mechanism for ‘steering at a distance’, and in particular as a top-down managerial device for reforming the structure and functioning of both the higher education sector and individual universities, along the lines envisaged by the policy elite” (Vidovich 2002, p. 397). The control of the tertiary sector by the government in the managerial framework runs counter to the configuration of universities as autonomous commercial businesses. This can be explained by two countertrends in tension with each other found in NPM whereby universities’ autonomy is neutralized by government using “systems of output-based planning, competition for resources, tied funding, performance reporting, accountability and audit” (Marginson 2010, p. 6). Therefore, the original aim of decentralization of universities has in fact turned into a ‘controlled decontrol’.

Moreover, the discourse of quality assurance has acquired a strong commercial emphasis particularly since 1996, particularly after Minister Kemp (1999), who turned quality assurance from a management to a marketing device. In Vidovich's (2002) own words: "Into the 2000s, the Government is still a major quality assurance player and it has positioned itself as a market manager for HE" (p. 406). In fact, discourse played a role as "Reid's (2009) study of quality assurance in Australian HE shows how the social practice is shaped by discursive practice" since managers' discourse derives from "texts issued by a central agency which thereby became a central authority 'disciplining' universities to follow a particular approach" (Cuthbert 2011, p. 134). The spread of managerialism in HE is considered the consequence of "increasing dissemination of NPM rhetoric and narratives" (Bleiklie et al 2011, p. 1).

In this scenario, Harman (1998) cautions that

Evaluation and reviews are therefore not just technical mechanisms for accountability and improvement but can be used and are used by different political actors associated with higher education as important policy instruments:

[...] But whatever the objectives, we need to be conscious of political agendas and motives in any programme of evaluation and reviews, particularly at the national or higher education system level. Of the range of policy instruments available, quality assurance reviews can be particularly powerful policy levers in the hands of ministers and bureaucrats." (p. 347).

In fact, McTaggart et al (1991) point out that the attributes of evaluation in managerialism raise dangers about fairness and even doubtful effectiveness, the very principle managerialism claims to be working for. They also defend that the vagaries of political decision-making risk remaining more influential in determining the fate of programs, than the findings of careful evaluation studies. In the same vein, the criterion used to do evaluation in universities with an NPM methodology runs the risk of becoming too focused on economic exercises since NPM's main concern is "achieving bureaucratic efficiency, while

increasingly aiming to reduce public run institutions' dependence on State funds." (Martínez Marco and Pérez de León 2019, p. 94).

1.2.3.3 Summary and Research Questions

The business-oriented management system implanted in the Australian Tertiary sector as a result of the Dawkins reforms thus presents a number of risks. The character of Quality Assurance requirements fostered by the Australian governments since the Dawkins era have evolved to become a policy instrument to control universities which paradoxically neutralizes their autonomy as commercial enterprises. In this context, reviews on behalf of Quality assurance (Q.A.) processes emulate the same approach within universities and have become either or both a policy and marketing check which allows managers to monitor whether programs, departments or schools are aligned with policy priorities.

Reviews started as a top-down device for reform to become a small-scale top-down management exercise. Increasingly, they have acquired the character of instruments of compliance to ensure agendas are implemented to do primarily with economic efficiency. This implies a focus on two main procedures: one, that marketing and other similar activities are put in place for programs, departments or schools to generate revenue and become self-sustainable; and two, and as a related consequence, a focus on keeping resources for educational affairs at a minimum. Thus, the potential educational benefits of regulated Q.A. reviews in the Australian HE system with an NPM framework are seriously being undermined, since educational improvement seems to be secondary to resource efficiency. As long as the priority in the NPM agenda focuses on turning programs, departments and schools into small private enterprises responsible for their own funding to keep afloat, reviews will do little to improve the educational side of programs which is what mostly

affect students, an approach radically different to the use of evaluations in other international educational systems, for example in language education in the United States.

Moreover, the accountability required of these units in Australian universities with an impetus to implement the efficiency agenda, have created an unstable and even a fearful atmosphere for all employees. As it is the case in any institution, power structures with no checks and balances pose a serious risk since they can lead to a climate of power manipulation for self-serving purposes.

In recent years, with corporate management well established in Australia, evaluation discussions and proposals revolve around disseminating and educating about an expanded understanding of both accountability and evaluation without necessarily calling for participatory forms of evaluation. As an example, Owen (2004) considers different meanings of accountability and stresses the concepts of fairness and attribution, and the need to avoid potential negative consequences which are “heightened when a review is undertaken” (Owen 2004, p. 3). Particularly, departing from the fundamental assumption that the right of the public to know is at the heart of any democratic structure, Owen highlights ten characteristics of current accountability systems. Among them is the lack of evaluation expertise particularly at not-for-profit agencies, which impacts on quality of evidence collected and disseminated, and a need to divert program development resources to evaluation. The other nine characteristics of evaluation embedded in accountability systems are (1) Increased community interest in public expenditure’s achievements and attention to the needs of consumers; (2) Focus of responsibility moves from the centre to departments or agencies with the expectation that reporting will be provided back to the centre; (3) Differing accountability imperatives according to hierarchical provisions; (4) Likelihood that not all stakeholder groups’ needs will be met in evidence-based systems; (5)

Evaluation incorporated into accountability arrangements; (6) Evaluation equated incorrectly with development and application of performance information; (7) Widespread concern that accountability information is not being used; (8) Inability of outcome focused frameworks to adequately tap the range of outcomes of a given intervention and to provide practitioners with information that leads to improvements; (9) A narrow view of what counts as evidence.

The focus on the economic efficiency agenda seems to have permeated deeply into the power structures and its implementation seems to be the key priority. Managing evaluation-informed processes with evaluation expertise can help raising the quality of the evaluation process itself, thus ensuring good practice so that not only policy and accountability functions are driving evaluation. The following study will attempt to fill this gap by demonstrating the role of evaluation expertise, and the usefulness of metaevaluation in exposing the (in)validity of evaluation practices, in the hope it will be incorporated more widely.

Based on the literature review presented in this chapter, the key research questions informing this study are:

1. Is metaevaluation an effective method to find out the character of the review processes of language departments across Australia?
2. What is the character of the evaluations – is their focus mostly managerial or educative? What are the implications of this for the theory and practice of evaluation and metaevaluation?
3. How can the key findings of this metaevaluation help guide future LPEs align with good evaluation practices within a strongly ideological policy context?¹²

Chapter 2. Conducting Metaevaluation

Chapter 2. Conducting Metaevaluation

2.1. Introduction

Metaevaluation constitutes the theoretical framework of this thesis, both as a concept and as a methodology. It serves several objectives:

- a) “metaevaluating evaluation processes” in seven case studies,
- b) conducting Document Analysis using an Evaluation Report Checklist as a guideline to explore contents of evaluation reports supported and contextualized by interview data from report participants, and
- c) reflecting upon the challenges that have emerged in the planning and the conduct of this metaevaluation.

In this chapter, the methodological framework for this multilayered metaevaluation process is presented: both the rationale for choosing a metaevaluation methodology which specifically responds to the main aim of this study, with the conceptual framework supporting it; and a justification of how the methodologies that have been selected to carry out the analysis of data fit into the general framework which overall constitutes a multiple method approach.

2.2. A Multiple Research Methodology Approach For Metaevaluation

In program evaluation, as well as in other social science and educational areas, the use of a variety of methodologies and types of data is encouraged as long as the context allows it (see Patton 2002, Bowen 2009, Johnson 1992, Mark, Henry and Julnes 1998). This diversity enables the evaluator to study a program from various angles, providing opportunities to gather multiple perspectives which can contribute to the most comprehensive understanding of the evaluand possible. Furthermore, the diversity of data

and methods, a process known as *triangulation*, strengthens the validity of the evaluation and therefore also in metaevaluation. Triangulation, according to Alkin (2011), “is a hallmark of qualitative analysis since significance testing is impossible. The claims that you make will have greater validity when they are based on consistent patterns across multiple data sources.” (p. 184).

In this metaevaluation study of Australian University language programs, given the importance attributed in evaluation theory to the idea that the evaluation should be conducted in a principled and systematic manner, and with the characteristic rationality and purposefulness of ‘evaluative thinking’, validity is addressed by adopting a well-defined multiple approach to methodology. This will contribute to ensuring a variety of perspectives and a quality metaevaluation process thus increasing the credibility of its findings. The role of the evaluator is also considered an important validity issue in terms of its independence in the evaluation literature. In this metaevaluation, the role, the values and interests of the metaevaluator in conducting the metaevaluation are made explicit, thus acknowledging any potential bias. Finally, validity is also strengthened by understanding the relevance of being responsive to the ethical guidelines required by the host university and by the needs and values of the participants in the study in addition to other contingencies that may emerge in the conduct of the metaevaluation.

In metaevaluation theory, efforts for prescribing procedures have emphasized not so much recommending theoretical orientations or models, but “central issues of questions and criteria” on the one hand and methodologies on the other (Wingate 2010). Hence, over the years, a variety of suggestions have been presented for how metaevaluators should review evaluations and what the focus should be which are discussed in the sections below.

2.2.1. The Measuring Instruments: Metaevaluation Criteria and Metaevaluation Quality

When considering the main role of a metaevaluation to judge the quality of an evaluation, challenges arise on agreeing both on criteria to be used and a definition of *quality*. As far as criteria are concerned, the use of standards not only contributes to professionalizing evaluation practice but “aids in educating evaluation consumers about what they can and should expect from evaluation” (Wingate 2010, p. 771). In Stake and Schwandt’s (2006) view, comparison is fundamental whether it is to “formal standards, silent expectations, to randomly selected groups, or to ways things used to be”. Various versions of standards, checklists and other quality criteria have been formulated for evaluators to ensure professional and high-quality evaluation processes. One of the most influential ones have been *The Program Evaluation Standards* by the U.S. Joint Committee on Standards for Educational Evaluation which comprises members from different organizations. The aim of *The Standards* is to provide guidance before, during, and after the evaluative exercise and the most updated publication of these standards emphasizes the following five essential attributes: (1) *utility* (to ensure that evaluation serves practical information needs); (2) *feasibility* (to ensure that evaluations are realistic and prudent); (3) *propriety* (to ensure evaluations are conducted legally and ethically); (4) *accuracy* (to ensure evaluations convey technically adequate information); and (5) *metaevaluation* (to conduct evaluations of evaluations) (Mertens and Wilson 2012, p. 23). However, the professionalization of evaluation and the persistent focus on methodological inquiry has contributed to the perception of evaluation as a technical activity (Bustelo 2003), a concept that is contested by some evaluators who believe that “Evaluation cannot be reduced to a simple procedure in which one follows a finite series of rigidly prescribed steps, nor is it

uncomplicated conceptually.” (Jarvis and Adams 1979, p. 4). This idea is an important one that contributes to the rationale of this thesis in the sense that bringing to the fore its conclusions for discussion among different audiences may provide a forum to begin to understand what is happening with evaluation practices and why it is happening according to research findings.

Methodological quality began as a concern in evaluation because of the use of evaluation by non-experts but according to Patton (2002) it is not as important as long as “methods and measures are adapted to the varying situations” and so evaluators focused on methodology because, evaluators thought, it helped maintain an image of power and prestige for the profession while knowing that “decisions on methods are rarely, if ever, purely technical” (p. 136). However, evaluators anticipate some risks in relying on standards. For example, Cronbach and associates (1980), in a period when, as discussed in the literature review, evaluation was in need of reform, while acknowledging the need for continuous review of evaluations, identify a danger deriving from too strong oversight:

Standards frozen prematurely could lock evaluation into unproductive practices and impede the intellectual growth of the enterprise. The best safeguard is *multiple, independent sources of criticism* [emphasis added]. Differences in values and orientations or overseers connected to various institutions are not only inevitable but are to be cherished. As long as the conflicting perceptions are expressed with equal force, the tensions will promote understanding and discourage any tendency to fit all evaluations into one mold. (p. 364)

More recently, further challenges have been raised to applying standards due to their strict nature. For example, even though the application of standards ensures a comprehensive metaevaluation, Scriven (2009) cautions that a ritualistic application may lead to missing important points. Cooksy and Caracelli (2009) note that when conducting metaevaluations “in the spirit of The Program Evaluation Standards” (p. 2) the scope of use is reduced or limited to evaluations of single evaluations in contrast with evaluations of

multiple evaluations, as is the case in this thesis, where the organizational capability for evaluation of a group or an organization can potentially be more comprehensively assessed.

When the essence of metaevaluation is to evaluate evaluations in order to ensure and improve their quality, there emerges the need to clarify what quality means. In fact, the concept of evaluation quality is an essential one in evaluation because of its direct association with the instrumental use of findings (Nevo 2009). Similarly, in this thesis I am using metaevaluation to explore policy and its impact, situating the centrality of New Public Management (NPM) to the thesis reviews, discovering an additional purpose for metaevaluation.

Since defining and discerning quality has been identified as a major challenge, the lack of agreement on a universal definition of quality may be identified as a possible pitfall of metaevaluation. In educational evaluation, the meaning and assessment of evaluation quality is currently still regarded as an urgent issue to resolve, particularly in the area of assessment of qualitative evaluative practice (Datta 2006).

In Stake and Schwandt's (2006) view, the conception of quality is bound to be influenced by the values of observers as well as the evaluator. In other words "in the best circumstances, the activities that constitute judgment of the quality of projects, programs, even entire evaluations, [...] are "political" in the sense that the judgment is a function of a commonality that can be exercised only by citizens interacting with one another in the context of mutual deliberation or decision" (Barber 1988, p. 200). The suggestion put forth by Stake and Schwandt (2006) thus consists of the study of both quality as measured as well as quality as experienced, as they state in the following quote:

Like other ways of assessing quality, metaevaluations too draw on conceptions of quality as measured (e.g. checklists and scales, audits of procedures, and data) as well as quality as experience (e.g. the judgment, grounded in practical experience, of

peer evaluators and stakeholders). Procedures for such metaevaluations are hardly systematized or standardized in most cases, but they already are part of the best evaluation studies (p. 416).

According to Cooksy and Caracelli (2005) quality criteria must then be “tailored to the purpose of the metaevaluation and to the culture and sensibilities of the metaevaluation’s stakeholders” (p. 35). In this thesis, the selection of quality criteria is primarily linked to the general aim of the metaevaluation. This method of analysis, having an explorative character, uses fundamental evaluation issues, including context, to research reviews in a comprehensive way. In fact, data and measurement instruments such as checklists and stakeholder perspectives on the practical experience of reviews will contribute to addressing Stake and Schwandt (2006)’s concept of “quality as measured” and “quality as experienced”.

2.2.2. The Metaevaluator

The actual control and power assigned to the metaevaluator is a key concern in metaevaluation and it is most commonly raised through evaluation practice. The reviewing of an evaluation has been conceived and claimed to be a professional activity, hence, the emphasis on the role of evaluators. For Cronbach and associates (1980), the scholar in charge of the review should never “speak for” the evaluators, which they claim is “indispensable for the health of the evaluation enterprise” (p. 363). That is, to keep the integrity of the metaevaluation and to avoid any validity threat, apart from adequate training and experience, the best potential profile of a metaevaluator may be a ‘critical friend’ or even better a ‘smart enemy’ or ‘tough rival’ (Scriven 2009) and not dishonest metaevaluators who may seek to detract primary evaluators for their own interests such as obtaining future contracts (Wingate 2010).

The *utility* standard from *The Standards* promotes the idea that the evaluator should be competent but also trustworthy. In this respect, Wingates (2010) has identified that the metaevaluation may engender a lack of trust in primary evaluators, especially if metaevaluators feel pressure to be critical about the conclusions reached in primary evaluations which may lead to misunderstand the function of metaevaluation. Scriven (2009) verifies that metaevaluation for some people may look like “a confession of incompetence” (p. iv). Hence, it should be made clear to stakeholders that “metaevaluation is part and parcel of sound evaluation practice and is not intended to discredit a primary evaluation, but rather to enhance its utility and /or credibility” (Wingate 2010, p. 771).

When the lack of practical experience in the field of evaluation may be considered as a weakness in the study, other arguments may compensate it. In this metaevaluation, the researcher acknowledges lacking training and experience in the conduct of evaluations or metaevaluations. However, insights on evaluation are contributed via the formal study and research on evaluation theory and accounts of real evaluations. Also, participation in formal whole evaluations of language programs in US universities as a member of staff, in formal program embedded evaluation activities such as regular class observations, and informal evaluation activities carried out by herself such as mid-term student evaluations, constitute useful practical knowledge to contribute a variety of perspectives. It must be said that it was the involvement in these evaluative practical experiences what sparked her initial interest in PE in the Australian higher education context. In this metaevaluation, having researchers from outside the evaluation profession conducting metaevaluations may be seen as beneficial, since their contribution may be perceived as a fresh outlook into the intricacies of particular evaluation practices in specialized educational fields such as Applied Linguistics; it may contribute to open an interdisciplinary dialogue between those involved

in the evaluation field and those in applied linguistics genuinely interested in LPE; and to illuminate not only what characterization the evaluation practices under study have but also begin to understand why. To address the potential political challenges mentioned above, it is thus important that evaluators make explicit their values. In this metaevaluation, the metaevaluator has:

- (1) a research interest as mentioned above to provide a professional, high quality, comprehensive metaevaluation that takes into consideration evaluation theory but is also interested in the perspectives of those on the ground and actual evaluation experiences and the results of evaluation;
- (2) an educative interest on the role of evaluation as an efficient tool for self-knowledge
- (3) a practical interest for the potential of metaevaluation to contribute to help create a PE framework for languages programs in universities.

It is the view of the metaevaluator in this thesis - based on her professional and educational experience - that language departments in universities should have LPE and LPM processes integrated in their operations because

- (1) it offers them tools for growth;
- (2) it can help raise standards in language education;
- (3) it offers opportunities for professional development;
- (4) when SLO (student learning outcomes) are made part of evaluation processes, findings may be used for both improvement and accountability purposes.

However, the analysis of data in this thesis is done in a spirit of embracing all kinds of findings, expected or unexpected, in order to provide realistic explanations.

2.2.3. Metaevaluation Focus

The persistent focus on evaluations of reviews “of, by, for and even about evaluators” is questioned because it implies that the perspective of users is hardly ever included (Cronbach and associates 1980, p. 359). Furthermore, according to Raizen and Rossi (1981) when metaevaluation speaks more directly to the uncertainties of diverse constituencies it has the potential to increase evaluation use. For Cronbach and associates (1980), a focus solely on the performance of evaluators will not provide the entire picture of an evaluation. Hence, the review of an evaluation should:

[...] consider whether the task set for a particular evaluation made good use of the energies of evaluators, commenting even on the wisdom of setting up an evaluation of that program at that time. That is to say, critics should review what commissioners do as well as what evaluators do. A contract that overspecifies the work, or a process of transmission that muffles the findings, is an abuse. Abuses deserve exposure. Criticisms should appreciate any difficulties that evaluator and sponsor encountered and overcame; beyond that, it is proper to complain about difficulties unresolved. But a review is likely to concentrate on alleged faults. Reviewers want their keenness to be admired, and sometimes they aim to tear down findings they dislike. We want evaluations to be reviewed in a collegial rather than adversary spirit. (p. 359)

Additionally, acknowledging that evaluations are ‘politically charged’, Wingate (2010) and Shadish, Cook and Leviston (1991) argue that a focus on critiquing methodologies of an evaluation may lead to missing the social and political effects of the evaluation which may be positive or negative.

The scope of focus in this metaevaluation intends to be as broad as possible given its exploratory character. Hence, it not only examines the performance of evaluators and commissioners and what methodologies guide the evaluations, but it studies the evaluations using fundamental evaluation issues to ensure most aspects of the evaluation are taken into account. The aim is to expose both successes as well as abuses in the

evaluative practices under research. But what methodologies are available to conduct metaevaluation?

2.3. An Emergent Realist Evaluation (ERE) Model for Conducting Metaevaluation

The process of reviewing an evaluation has been compared sometimes to that of the review of scholarly work by peers (Scriven 2009), and to the review of a book or even a law where alternative explanations are suggested by other professionals (Cronbach and associates 1980). As far as specific methodologies for conducting metaevaluations, several models have been proposed. In the past, in order for the reviewing of an evaluation to have an open character, one suggestion was to review evaluations at pre-arranged sessions in professional meetings or symposiums where preliminary designs and reports could be distributed for discussion and critique, or by having organizations commission panels to review evaluations, which, having a bearing on educational policy, could help make reviews “sophisticated and responsible” (Cronbach and associates 1980, p. 363).

Cook and Gruder (1978) suggest a framework meant to encourage the, then, *new* systematic evaluation research. Identified within the positivistic tradition, their main focus was the quality of methodological procedures in general, and particularly on their technical quality, which they verified as deficient in the four studies of evaluations that they had conducted. A different proposal to appraising evaluation work through metaevaluation is the one by Gowing and Millman (1978), which consists of three approaches, namely, (1) standards for evaluation; (2) a checklist of questions for evaluation; and (3) the QUEMAC framework (Question, Even-Object, Method, Answer, and Concept) which consists of a series of questions seeking to reveal the logical structure of the evaluation. Several additional checklists have been composed to guide the conduct of metaevaluations. One of

such is Stufflebeam's (1999) checklist based on *The Standards for Program Evaluation*, and Scriven's (2007) Key Evaluation Checklist to conduct evaluations suggests metaevaluators applying professional standards such as the ones by the Joint Committee when they arrive at the metaevaluation checkpoint.

Stufflebeam (2001b) and Stufflebeam and Shinkfield (2007) also contribute with a framework of eleven generic tasks that metaevaluators can use for heuristically planning their metaevaluations. The procedures for using these tasks are context-dependent and therefore it is the metaevaluator's job to decide which methods will be more effective within each particular context.

Metaevaluation can also be completed using *The Program Evaluation Standards* (Joint Committee on Standards for Educational Evaluation 1994) themselves. As mentioned, there is a metaevaluation standard with which an evaluation may be rated according to the extent to which different aspects of the metaevaluation standard were addressed, partially addressed, not addressed or not applicable (Wingate 2010). In fact, the standards alluded to by Gowing and Millman (1978) anticipate the creation shortly afterwards of the standards produced by the association *Joint Committee on Standards for Educational Evaluation* (1981) which is now in its third edition (2011).

In the higher education context, Saunders et al (2011) bring forth a metaevaluative tool named RUFDATA, an acronym that stands for Reasons, Uses, Focus, Data, Audience, Timing, and Agency originally conceived to plan an evaluation and 'get it off the ground' (Saunders 2011, p. 17). This tool is used, though not comprehensively, by evaluators or evaluation participants at various international higher education institutions to reflect on the real evaluations in which they got involved and then retrospectively write an account about their insights. Subsequently, Saunders et al analyze these accounts again using

RUFDATA in order to see whether different kinds of practices within one domain (systemic, programmatic, institutional or self-evaluative) reflect similarities and differences and what can be learned in order to strengthen and improve them.

Regardless of which checklists or frameworks are selected, Scriven (2009) cautions against “scientific fundamentalism” and advises the metaevaluator to use frameworks flexibly so that if ‘piecemeal work’ is desirable, he or she is still able to think about “what the report is not doing, what it might be doing, or what it is doing viewed from a holistic point of view” (Scriven 2009, p. v). Keeping Scriven’s suggestion in mind and the conclusions about theoretical orientations and methodologies to conduct evaluation reached in evaluation theory, this metaevaluation uses an ERE framework which, while following the procedures in Miron (2004), is conceptually similar to that proposed by Mark (2008). While Mark’s framework was originally created for conducting evaluation, it is replicated in Mark’s taxonomy of subjects and inquiry modes to conduct research on evaluation (2008), in which metaevaluation is presented as a candidate research approach and methodology. This framework conceptualizes evaluation in a holistic way, thus reflecting the trend in evaluation towards more inclusive evaluation models (Mark, Henry and Julnes 2000).

Mark claims that metaevaluation can involve different kinds of research and for metaevaluation to properly belong to a category of research studies, it cannot consist merely of a review of the judgments carried out by an expert. Rather, the types of research studies that Mark brands under metaevaluation are of two kinds, one of which suits the aim of this metaevaluation, namely, “a comparative meta-analysis across several evaluations and examination of the relationship between some aspect of evaluation context and the kinds of evaluation activities that were carried out” (Mark 2008, p. 128). In the past, the focus of evaluation research was meta-analysis or research synthesis with a focus mainly on

numerical data, but Cronbach (1982) warns the reader against the use of meta-analysis, because it “can degenerate into scientism when the several studies do not aim at a single, well-defined target” (p. 310). In this thesis, Meta-analysis should be understood as a synonym of metaevaluation.

Set within the ERE model, evaluation methods are discussed as technologies that humans have evolved to supplement natural sensemaking. Methods, in this theory, are mostly concerned with revealing the underlying mechanisms that are at play because those will help understand the program the most but at the same time it is recognized that the explanation of underlying mechanisms might not represent the only means. Thus, ERE puts an emphasis on contextual factors as the basis that will likely dictate the extent to which such explanations need to be part of evaluation planning.

Also, and in comparison with other types of realist evaluation frameworks, heightened attention is given in this model to the potential fallibility of methods which therefore need to be critically assessed in order to learn about their suitability and limitations. For this reason, ERE evaluators will best respond to the needs of a particular evaluation by using multiple methods. Another significant aspect of this new paradigm is utilization, understood as the action generated by evaluation conclusions. If the purpose of evaluation is to ameliorate society, then the effect of evaluation has to be felt across the policy-setting community¹. ERE emphasizes values-probing to the extent that, for example, different groups of stakeholders may hold a variety of perspectives about what is *important information* about their program. These diverse views may affect the focus of the evaluation because depending on the type of information that is prioritized, the evaluation may strengthen one view over another. Hence, the study of values is integrated into the ERE theory.

In sum, the integrative framework offered in ERE amplifies the vision of evaluation. On one hand, it moves away from what ERE theorists consider artificial paradigmatic dichotomies such as the 'paradigm war' and on the other it provides a framework particularly emphasizing the inclusion and explicit study of values as an additional component to sensemaking.

2.3.1. A Taxonomy of Subjects of Inquiry and Inquiry Modes for Metaevaluating LPEs in Australian HE

The processes that from a realistic perspective provide an explanation of what happens in the world are akin to the processes of description, classification and causal analysis used traditionally to conduct evaluations (Mark, Henry and Julnes 1999). However, in the neorealist integrative framework of ERE values inquiry is advocated as an additional category (Mark, Henry and Julnes 1998). These four inquiry modes, thus, are suggested together with a classification of evaluation purposes to guide evaluators' work.

This sophisticated model offers a comprehensive and inclusive theoretical as well as methodological basis for metaevaluation which involves the use of several methodologies to conduct a comparative analysis of the evaluation context of a series of evaluations within a system. In the section below a description of the taxonomy of subjects of inquiry and inquiry modes and its application in the analysis of data is presented. However, since in this framework, subjects and modes are selected according to the purpose of the evaluation to ensure a useful choice, a purpose or purposes for the metaevaluation in this thesis need to be first articulated.

2.3.1.1. Purpose(s) of Metaevaluation

As in educational evaluation, formative and summative purposes can drive, too, the use of metaevaluations (Stufflebeam 1975). A summative metaevaluation assesses the primary evaluation after it is completed and provides information for accountability purposes to “confirm (or disconfirm) that an evaluation’s conclusions (and recommendations, if present) were sound and justified” (Wingate 2010, pg. 766). That is, this kind of metaevaluation seeks to examine the extent to which the primary evaluation addresses the requirements of utility, feasibility, propriety and accuracy as defined in *The Standards for Program Evaluation* (Stufflebeam and Shinkfield 2007). One significant benefit identified in using summative metaevaluation is that decisions based on faulty information may be prevented when pitfalls in the primary evaluation are found (Wingate 2010).

Scriven (1998) advocates specifically formative metaevaluation, as it allows the quality of an evaluation to be assessed. The aim of metaevaluation for formative purposes is to give feedback “at different stages of the evaluation’s life, such as in the planning stage, in the middle, and at the end” for improvement (Scriven 1998, p. 38). Stufflebeam and Shinkfield (2007) claim that the feedback should focus mainly on the level of adherence to the original evaluation design. Also, since the evaluation’s inadequacies are addressed, it helps “preventing the determination and dissemination of invalid conclusions and increasing the primary evaluation’s utility and cost-effectiveness” (Wingate 2010, pg. 766) and it may even be included in the final report (Shadish et al 1991). Formative purposes for metaevaluation are also claimed to increase the confidence¹³ of those involved in the evaluation (Mertens and Wilson 2012). However, metaevaluation practitioners may use summative metaevaluation formatively “by taking what was learned in a summative

metaevaluation and applying those lessons in their subsequent evaluation experiences” (Wingate 2010, p. 766).

By contrast, several other authors have claimed the multipurpose nature of metaevaluation beyond summative/formative purposes. First, a metaevaluation of a self-evaluation conducted at the Universidad de Almería, Spain, enabled researchers to verify that, besides accountability and enhancement purposes, metaevaluation helps guide “strategic organizational change and it legitimizes evaluation systems” (Reboloso, Fernández-Rodríguez and Cantón 2009, p. 16). Second, the findings from a metaevaluation of a series of evaluations of gender-equality policy led researchers to conclude that “important conclusions and guidance to improve the evaluation function in a specific field can be drawn from the systematic analysis and assessment of evaluation practices, their strengths and weaknesses and their final utility” (Bustello 2003, p. 400). Third, as Saunders (2011) commented above, metaevaluation is used to formulate an evaluation approach that is based on practice and offers guidance for conducting evaluation in the specific context of higher education. In particular, a set of maxims and implications for stakeholders, evaluators and users of evaluation in general are provided.

Wingate (2010) provides information on two more purposes found in the research literature regarding metaevaluation, namely, (a) a purpose which focuses on the synthesis of findings across evaluations of a particular program or set of programs in order to understand more deeply how effective programs are; (b) a research purpose for metaevaluation which seeks to identify whether the following hypotheses suggested in evaluation theory are addressed in the evaluation:

1. Use and application of evaluation standards and principles to improve an evaluation’s quality and/or ethical problems.

2. Stakeholder involvement to enhance use of evaluation findings.
3. Participatory and collaborative approaches used for capacity building to enhance program effectiveness and increase evaluation use.
4. Employment of a variety of reporting strategies besides the written report, applied to diverse audiences to increase use of findings (Wingate 2010, p. 766).

In sum, alternative purposes have been used in metaevaluation practice which go beyond purely formative and summative ones. In this thesis, the purpose of the metaevaluation, thus, appeals to the versatility of purposes and includes multiple aims.

Since the metaevaluation undertaken is part of a doctoral dissertation, it has an important research function. It aims at filling in a research gap in applied linguistics to address the need to ensure that “scholarly traditions of metaevaluation are fostered” (Norris and Watanabe 2013, p. 5).

To fill this gap, it seeks to contribute to building on empirical evidence about evaluation in general and in particular of language program evaluation which in turn helps disseminate examples of evaluation practice and learn from its realities and, particularly, about (1) the importance of assessing contextual elements in the complex higher education setting; and (2) the potential usefulness of evaluation when conceived not as a policy instrument, not as an exercise for compliance but as an educational tool for improvement.

Third, according to metaevaluation theory, this metaevaluation is characterized as summative since it is conducted after the evaluations have taken place. However, it aims to use its findings to propose recommendations for conducting language program evaluations in university departments based on the learning derived from the findings, hence making use of summative information formatively. Furthermore, a formative metaevaluation option which involves an on-site observation was disregarded early in the study since the pilot

study conducted prior to this thesis had revealed a general overt confidential character of reviews. Moreover, to provide a broad picture of the character of evaluation practice in Australian universities language departments, the number of review cases that could virtually be analyzed via interviews and reports would considerably be higher than if the on-site option was chosen because these evaluation processes “occur rarely”, as participant (S) confirms in one of the interviews conducted for this thesis.

Additionally, as in evaluation in general, the internal and external distinction is also highlighted in metaevaluation theory. However, in contrast with the traditional association of internal evaluation with formative purposes and summative ones with external evaluation, while internal metaevaluation is associated with the changes that are applied to improve the quality of the evaluation as the evaluation goes on, generally using *The Standards*, external metaevaluations are conducted by an external party and can be used for summative or formative purposes (Fitzpatrick et al 2011).

In this thesis, while the metaevaluation is fundamentally external, it may be considered to have some degree of internality. The field to which this doctoral thesis is affiliated is applied linguistics and the background of the researcher is language teaching and learning supported by years of education and professional experience in languages education which make it to an extent a *peer metareview*. Hence, to carry out a cross-examination of a series of language program review processes in Australian higher education this metaevaluation orients to subjects of inquiry which evaluators in the domain of languages programs in higher education have found essential to conducting meaningful evaluations which involves the investigation of

- (a) The extent to which the evaluations under study address these issues (namely, the role of the evaluator, the involvement of teachers in the evaluation, the variety

of reporting strategies, the variety of data collection methods, triangulation of data sources and methods)¹⁴

104(b) The character of these evaluations, formative or summative

(c) The organizational/policy context of universities

(d) The group of values which create balance, or imbalance in the evaluation process.

These metaevaluation purposes are in fact interrelated in the sense that, for instance, the study of the context (purpose “c”) may reveal the values in the evaluation (purpose “d”) which in turn entails a thorough description of the issues that lead to successful language program evaluation processes (purpose “a”). This description, in turn, will shed light on the evaluation attributes that will help classify their character, thus addressing purpose (b). Accordingly, to carry out the analysis that addresses these purposes the use of inquiry modes entails values inquiry which can typically be combined with description and classification¹⁵ (Mark et al 1999), as will be the case in this study and which is next explained.

2.3.1.2. The Description Mode

Description deals with events or/and experiences which belong to a level of reality that is more directly detectable. According to the realist evaluation model, reality is comprised of stratified regular patterns of meaning which may be directly observable while others are not. The methods that are suggested for description are quantitative as well as qualitative. When qualitative methods are used, description will typically involve the complexity characteristic of social contexts, and they will provide a full description of, for

example, evaluation reports including evaluation activities, participants, goal, evaluation use, reporting, stakeholder involvement, evaluators, etc.

2.3.1.3. The Classification Mode

Classification is concerned with categories or structures that are not readily discernible and need to be discovered via the use of methodologies for they live in a level of reality that is just “beneath the surface” (Mark, Henry and Julnes 1999, p. 185). The types of methods used for classification in the framework include qualitative, quantitative and mixed methods. Once categories of “patterns of observable characteristics” (Mark et al 1999, p. 182) can be inferred and a description of the attributes of each is developed, a decision will follow about which evaluation case falls within one categorical group or another, which will help compare the attributes across evaluations with those of the categories.

2.3.1.4. Values-Inquiry

The study of values in ERE is carried out taking into consideration that reality is complex and stratified. The ERE evaluator sees values as part of embedded systems and therefore it is desired that the methods used in an evaluation support this stratified valuation so that representativeness is achieved. As an example, it is normally the case that evaluation reporting only reaches certain stakeholders but ERE argues for a broad distribution of information, in order to reflect this representativeness.

The methodologies that account for the complexity of a particular evaluation context aim to achieve a balance of individuals, a balance of values and to take into account contextual factors. Values balance can be accomplished by investigating the full range of

stakeholders. Methods, thus, have to account for contextual factors that affect program outcomes. But which values shall be studied? ERE proposes choosing to study contextual factors according to *value relevance* (Henry and Jules 1998) based on the concept of emergent processes, that is, the values to focus on are those with a connection with human needs rather than desires (Henry and Jules 1998).

So far, very few examples of values inquiry used for research in evaluation have been found and Mark (2008) invites researchers to experiment with the four inquiry modes in order to accumulate knowledge on evaluation practice, thus facilitating a way to compare, to classify and to synthesize the findings that result from systematic research on evaluation. But besides the inquiry modes, this framework also offers a taxonomy of subjects of inquiry to deal with evaluation studies.

2.3.1.5. Taxonomy of Subjects of Inquiry

Four broad subject categories are proposed in this methodology namely evaluation context, evaluation activities, evaluation consequences and professional issues into which subcategories of content areas are created and correspond in the main to fundamental evaluation issues.

A *domain specific* category is not included in Mark's taxonomy for the broadness of categories will "better contribute to *general* evaluation theory and practice" (Mark 2008, p. 118). However, since the focus in this study is on evaluations of languages programs in higher education, the issues which have been found important in conducting "good evaluation practice" in this domain becomes especially relevant in this metaevaluation.

It is now time to turn to the kind of data and the methods used for data gathering, and the analytical framework carried out in order to address the research questions which will help characterize LPE practices in the Australian Higher Education sector.

2.4. Metaevaluation: Evidence, Collection And Analytical Methodologies

To enhance the validity and credibility of findings in this metaevaluation, a variety of data was planned to be used as suggested in the evaluation literature. Furthermore, as the pilot metaevaluation conducted prior to this thesis indicated difficulty in obtaining reports¹⁶, it became clear that learning from reports would need to be complemented by other kinds of data such as interviews and that being responsive to circumstances as data collection was in process would be paramount. Therefore, the plan for data collection was designed carefully with an open and flexible attitude to be able to gather the necessary adequate amount that would allow to conduct this research project. In fact, in metaevaluation theory, reports are used as valuable data but since they “cannot tell the full story about the evaluation” (Wingate 2010, p. 769-770), additional data is regarded as necessary (e.g. interviews, focus groups, surveys, etc.). At the same time, it was understood the need to offer assurance to those granting access to documentation or to participate in interviews of total anonymity and respect about the information they would provide. To achieve that, in this metaevaluation references to particular languages, universities, and interviewees are made with a previously elaborated anonymous code and to refer to specific interviewees a default 3rd person masculine subject pronoun (“he”) is also used.

The research plan for this thesis went through a formal ethics process whereby the host university’s human research ethics committee required the researcher to provide

detailed information about the project and additionally to write a plain language statement stating the purpose of the research project and a consent form which were to be read and signed by potential participants. These two last documents seek to assure participants the research-only purposes of the study, the confidentiality of information, and the sharing of findings with them if so requested.

Additionally, in metaevaluation, the issue of collaboration is discussed in terms of the identification of different interest groups or political viability in order to prevent “derailment of the evaluation process, bias, or misuse of results” (Wingate 2010, p. 772) which may have negative effects on the primary evaluation and on the people who have been involved with it. This issue could be perceived as a correlate to the issue of engagement or participation in evaluation discussed in the literature review which is considered necessary to increase the credibility and/or the objectivity of the study. Likewise, Kemmis (1986) recommends that, as in evaluation, care must be taken to avoid the possibility that the metaevaluation serves the interests of a particular group over another. Therefore, the metaevaluation will need to be validated which, as discussed earlier, in this thesis is achieved by adopting a multiple method approach thus ensuring a variety of perspectives. Validation in metaevaluation, in Kemmis’ view, may be attained by judging the metaevaluation against the same question that is raised for the evaluation, that is, “how does it contribute to the improvement of the critical debate about the programme as a whole, for the whole community of programme participants and interested observers?” (Kemmis 1986, p. 117).

Therefore, in this thesis, the data corpus includes verbal data from interviews conducted with individuals who have been reviewers or who have been at the receiving ends of a review, or both, as well as review reports of a set of evaluations from different

institutions within the Australian higher education system. The analysis of both types of data will help fulfill the aim of this metaevaluation which is to conduct a comparative analysis in order to explore the character of evaluation practice in that system.

2.4.1. Review Reports

The limited access and restrictions experienced when collecting the reports for this study led to the devising of three strategies to, in the end, obtain three new whole reports plus one Terms of References which added to the two reports from the pilot study made a total of six review reports, arising from both routine cyclical evaluative practice but more frequently from special intervention, often in the context of financial reorganization. The first strategy involved sending electronically a formal and carefully written letter by the researcher and overseen by the main supervisor to the heads of schools of each of the so-called Group of Eight universities¹⁷. The selection of these universities and their operational homogeneity was thought to aid in making generalizations about the current state of evaluation practice of a prestigious group of universities within the Australian higher education sector. However, the difficulty in gaining access to reports caused the researcher to resort to other non-Go8 universities which resulted in just one particular case providing an additional report which made a total of seven (7) reports altogether.

Moreover, in the process of sending away the letter to heads of schools, a contingency emerged which added complexity to the process of data collection. The research conducted for the pilot study was presented at the second bi-annual National Colloquium of The Languages and Cultures Network of Australian Universities (LCNAU)¹⁸ and the advisory committee of this organization, after considering the relevance of PE for

the future of languages in Australia, agreed to officially endorse the project that is pursued in this thesis¹⁹. The members of the LCNAU committee perceived the project as potentially useful to raise awareness about program evaluation among the language community in Australia and to possibly use ensuing recommendations to create an evaluation framework for language departments to evaluate their own programs, thus contributing to foster a culture of ongoing evaluation. For this reason, the idea of using the LCNAU logo in the letters came up as a possibility for it would help to reinforce the formality of its contents and confidential assurances to those whom the reports were requested. However, because this possibility was only informally discussed with one supervisor and not formalized due to legitimate reasons that the researcher was unable to gather at the time, and since the main supervisor had already supervised the letter, the researcher made the decision to send the letter away, perhaps now in retrospect precipitously, in light of the long wait that was expected that the reports-collecting process would take. Unfortunately, the LCNAU advisory committee found this a disrespectful action and became singularly vexed. The researcher and the main supervisor of this study immediately responded by apologizing and reassuring the committee about the original well-intentioned reasons for using the LCNAU logo (See Appendix).

The slow response rate to the letter and the little success in accessing reports via the first strategy urged members of the supervisory committee of this thesis to use personal contacts to aid in the collection of reports and although there was some response, this strategy became also insufficient.

Consequently, a second round of letters was sent again but this time to the vice chancellors of the same universities via regular post mail and just with the logo of the host university from which the letter was being sent. This action was in fact prompted by the

advice of one interviewee, (T), who encouraged the researcher to send the letter to upper administration instead because, as he advised, once an evaluation ends “reports are the property of the institution that instigates the review” and therefore “it has to be treated as any other document which is confidential to the institution” which is an idea supported by another interviewee, (M) who states that “M126. [...] They are documents that are reports to the vice chancellor effectively or the dean about how a department is doing.” However, no success was attained through this strategy either. In conclusion, the collection process of reports was lengthy (started in November 2013 and the last response letter was received in March 2016), inconsistent and unsettling at times and therefore the question about the determination of the number of review reports needed was subjected to what documents university administrators allowed to have access to.

In contrast, finding participants to interview and their willingness to collaborate was much smoother, though some interesting reactions were also found. For example, when the researcher asked in person to an Australian language educator about recent program reviews, the immediate response was “Oh, language reviews, God forbid!” which shows a clear negative reaction to reviews. In the evaluation literature and in particular in LPE, the resistance to evaluation is considered a challenge. The *uneasiness* with evaluation is thought to be triggered by discipline-related issues, and by “unfortunate and inaccurate” misconceptions of evaluation. Thus, several references are made about certain areas of the language education enterprise which have put it in a vulnerable position which arise at the time of an evaluation (Jarvis and Adams 1979) and have led to a crisis of confidence (Jacobson 1982). For example, (1) considering languages a non-essential aspect of education; (2) having a low status in public acceptance and doubts about the quality of programs due perhaps to promising greater proficiency than can be taught (Jarvis and

Adams 1979) but also perhaps to the lack of available information such as consistent goals at national or state level for basic skills in second language education, or “the ultimate goals of language study” which “result in a core of educators who resist being held accountable for student achievement”; (3) reliable program focus; (4) standards to measure program implementation and student achievement; and (5) effectiveness of methodological approaches (Jacobson 1982, p. 286).

Another common perception of program reviews particularly for accountability or accreditation demands is that of a burden and in extreme cases as “pointless and trivial” (Brown 2012) which make program staff view evaluation “as a nuisance required by a remote and nebulous agency; it is an exercise to be done so that it can be forgotten for another ten years” (Jarvis and Adams 1979, p. 2). One of the most serious consequences of these practices is that

Decisions about foreign language programs and staff training are being formulated today by those outside the profession –the urgent first step in regaining control for the profession is to be able to provide valid data to the decision makers. (Jacobson 1982, p. 287)

In the same vein, according to Patton (2002), negative past evaluative experiences whereby evaluation has been done *to* people rather than *with* people, particularly, the traditional top-down approach used mainly in carrying out program evaluation for compliance has caused program practitioners to feel skeptical about the extent to which their voices will be heard and second whether research findings will in reality be put to use. Evaluation done in this fashion has contributed to narrow perceptions of it as an external endeavor with little meaning for program staff (Mackay 1988). Hence, “Stakeholders consider imposed program evaluation a threat rather than an opportunity for help and improvement” (Shawer and Alkahtani 2013, p. 20) and program evaluation is undertaken as

an end rather than as a means of “knowing oneself and taking action, support for faculty development, recognition of valued institutional practice, collaborative inquiry turning program review into valued work... improvement, and impetus for innovation and ownership of programs” (Byrnes 2006, p. 576).²⁰

2.4.2. Interviews

The interviews needed for this research were conducted while the report collection process was being carried through and are the subject matter of the next section. As Figure 1 below shows²¹, interviewees were selected on the basis of attaining as much representativeness as possible in order to gain a variety of views. In total, 13 individuals were interviewed, out of which six had participated as evaluators and five had been on the receiving end of a review (i.e. ‘reviewees’ in Figure 1). To further expand the knowledge base of experiences in reviews, the researcher decided to include two external perspectives. On the one hand, the perspective of an Australian professional evaluator affiliated to the Australian Evaluation Society with ample knowledge and experience in the evaluation profession and field but with no connection to languages, and on the other, a non-Australian language professional with experience as an evaluator in Australia.

These interviews were semi-structured and contained basic questions to guide the interviewee but were also aimed at letting interviewees digress towards the topics they thought were salient in their experiences in reviews. The length of the interviews depended on the interviewee’s availability but ranged from one to two hours and were either done in person, via skype or telephone. The questions had two main parts (Appendix 1)²². The first one contained questions in relation to the actual conduct of the reviews interviewees chose to speak about. The second part sought to find out interviewees’ perceptions about

program evaluation in general. However, as interviews were being conducted, the researcher, herself an experienced university modern languages teacher, modified the format of the questions for the sake of clarity. As hinted up above, interviewees in general all cooperated eagerly and provided generous accounts of their experiences in reviews and therefore gaining buy-in was unnecessary in collecting oral data for this research project.

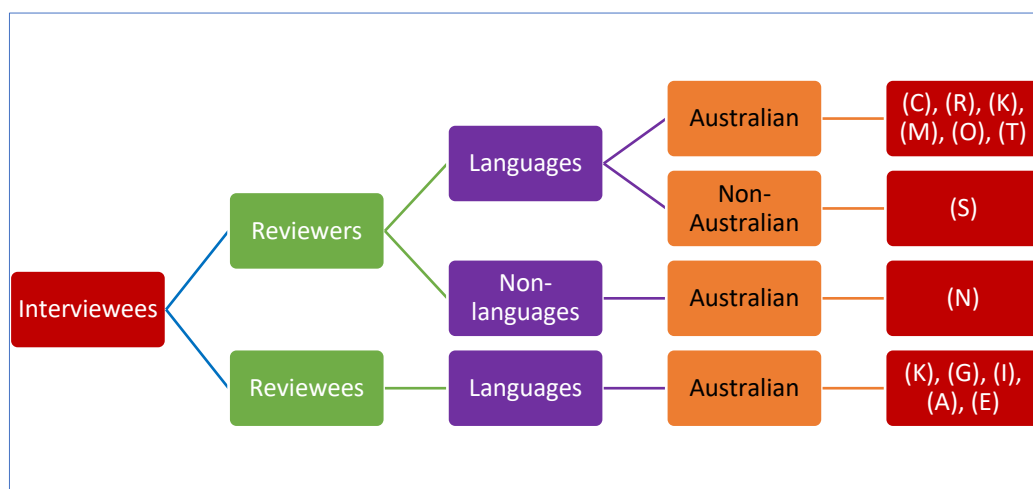


Figure 1. Interviewees

As the process of data collection in this thesis proves, metaevaluators do find contextual limitations as most evaluators do, (i.e. insufficient resources or unfavorable conditions), particularly to accommodate metaevaluations after each evaluation is conducted²³. However, as long as it is feasible, metaevaluation is encouraged since it “is an excellent basis for reassuring stakeholders and providing you [the evaluator] with feedback” (Alkin 2011, p. 250). Likewise, Scriven (2009) acknowledges the challenging endeavour that may be involved in conducting an entire metaevaluation in the real world while he also admits that even a partial metaevaluation is better than none. The methods chosen for the

analysis of data, the rationale for choosing them and the way they are applied in the metaevaluation are explained next.

2.4.3. Methodologies for Data Analysis

Since the focus of the analysis is on meaning, the methodologies selected to tackle it are qualitative although some quantitative data may be also integrated in the analysis (e.g. length of reports, etc.). Thus, on the one hand, reports are treated as a collection of ‘mini-cases’ which are studied intrinsically as individual case studies, and the data they contain is analyzed and then cross-examined following the guidelines of Document Analysis (DA) and an Evaluation Report Checklist (ERC)²⁴. The findings of these analyses are contextualized and fleshed out with material from the respondents in the interviews, which were transcribed and subject to a Thematic Analysis (TA). The transcription of data from the interviews involves the conversion of oral language to written text and the literature suggests that in fact, analysis starts with transcription since it is considered a “process of construction” (Hammersley 2012). Because a number of decisions, technical and interpretational, are involved in transcription, it is recommended that researchers state explicitly how transcription is done (Kvale 2009).

2.4.3.1. Transcription of Interview Data

In this metaevaluation, to provide answers to the research questions, the interest in analyzing data lies, as mentioned, in understanding meaning. Therefore, the transcription of oral language is based on its usefulness to facilitate the most accurate possible understanding of what interviewees mean by what they say. To this end, the following

decisions were made during the process of transcription to ensure a balanced interpretation of meaning.

The technical decisions that were made involved the formulation of a list of transcription conventions which was based on conventional symbols used in transcriptions of oral data and then adapted to the needs of the study. With regard to non-word elements, these were included in the transcriptions in case they would, in some significant manner, add meaning such as denoting expression of doubt, unfamiliarity with or little or lack of self-reflection on the topic.

The laying out of talk that is usually chosen for transcription is that of a playscript which is the one used in this thesis because it helps to visualize most clearly talking turns in the dialogue. Moreover, the labels for speakers was a changing process, from using the initials of participants' names for interviewees responses and the letter *Q* for the questions asked, but in the end for the sake of simplification, neutrality and clarity, the initials of names were changed to the letter *A, B, C* and so on for answers.

For the division of segments within the talk, grammatical sentences with punctuation such as commas and periods are used to help distinguish units with complete meaning and facilitate understanding of the ideas conveyed. Overlapping talk is also noted in the transcription because frequent overlap from a participant may denote uneasiness about the topic or may be feedback for improving interviewing style for the interviewer. This was in fact the case, and the transcription process became a learning process which taught the interviewer to minimize the interruption so as not to hinder the participant's train of thought.

Another point that is suggested in the literature to take into account is the length of transcription extracts that are to be used when reporting findings. In this case, the objective is to use sufficient data in an extract to exemplify in a clear way a particular finding.

Despite the lengthy process of transcription, in this study transcription was found very useful as an introductory analytical tool. The numerous times of listening to and writing verbal data did allow the researcher to form a view of both broad topics and sometimes detailed information and judgments by the interviewees which allowed for comparison with other interviewees' perspectives and also with data from reports. In the following paragraphs, a description is presented of the rest of the analytical methodologies used in this metaevaluation which include, besides document analysis (DA) and thematic analysis (TA) multiple Mini-Case studies (CS). The focus of this presentation is on the one hand their suitability for addressing the research questions of the study and their applicability to the analysis of the verbal data and reports. On the other hand, and more broadly, an explanation of the relationship between the analytical methods and the broader methodological framework of the study (Taxonomy of Subjects of Inquiry and Inquiry Modes) is also presented.

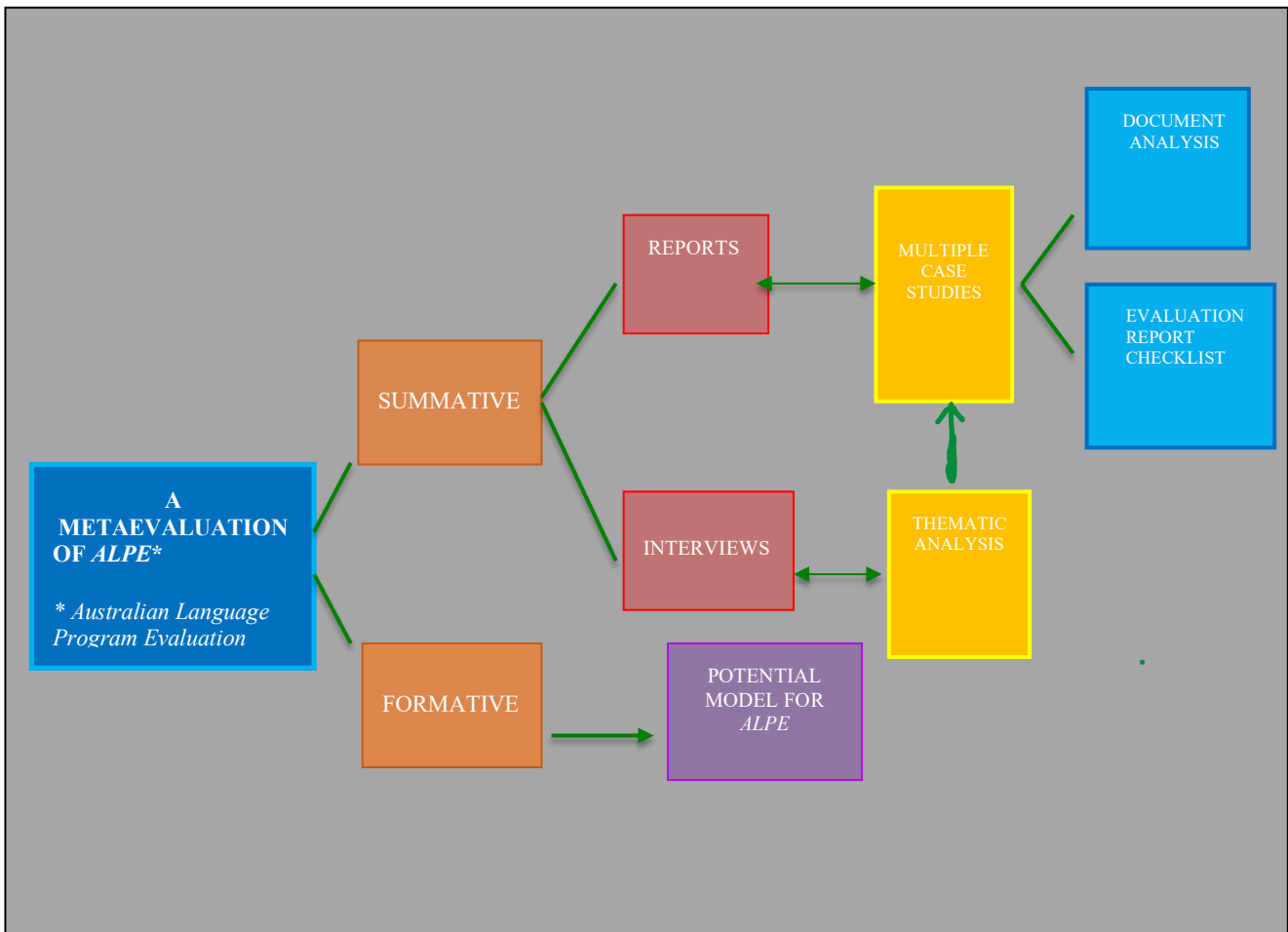


Figure 2. Multiple methodologies for a metaevaluation of Australian Language Program Evaluations (ALPE)

2.4.3.2. The Case Study (CS)

The research interest in learning about the character of a particular phenomenon such as LPE in Australian university departmental reviews, aligns with what the literature on Case Studies (CS) terms qualitative inquiry, considering that the CS is not so much a methodological choice as a choice of *what is to be studied* (Stake 2005, Flyvbjerg 2011). A prominent evaluator such as Cronbach in commenting about Stake’s (1986) metaevaluative case study which relied primarily on the same types of data as in this thesis remarks “Case studies of evaluations are welcome because they hold a mirror up to professional practice”,

(Cronbach, jacket note in Stake's 1986 book). The controversy about whether to use quantitative or qualitative methods for case studies has been treated in recent literature as having a secondary role in comparison to the primary relevance of determining the boundaries of the case (Flyvbjerg 2011), and that the strategies used for the analysis are aligned with the purpose of the study and the data collected (Johnson 1992). Whether quantitative, qualitative or both, a high-quality analysis in CSs is that which (1) identifies important variables, issues, or themes, (2) discovers how these patterns interrelate, (3) explains how these interrelationships influence the phenomena under study, and (4) offers fresh new insights (Johnson 1992, p. 90).

In this metaevaluation, the cases that are considered to qualify as such are those which are described thoroughly by interviewees, those which are in review reports and those which are both described by interviewees and are in a report format. These cases are then examined in order to establish a comparison and learn about the character of evaluation practice. While Stake, as an important contributor to case studies in evaluation, does not favor the use of comparison in case studies due to the risk of missing some important information about the singularity of the case, he however admits that "When there are multiple cases of intrinsic interest, then of course it can be useful to compare them" (Stake 2005, p. 458). He also adds that "Illustration of how a phenomenon occurs in the circumstances of several exemplars can provide valued and trustworthy knowledge." (Stake 2005, p. 459).

The study of these cases presents characteristics similar to what Stake (2005) calls an *instrumental collective case study* whereby various exemplars are examined since "it is believed that understanding them will lead to better understanding and perhaps better theorising about a still larger collection of cases." (p. 455). Since the case studies in this

thesis are contained within a series of documents (review reports) and verbal data (interview material), DA was selected as a second qualitative methodology to study particularly the reports, which is presented next.

2.4.3.3. Document Analysis (DA)

In program evaluation, DA is defined as a fact-gathering technique that allows the evaluator to select which facts mean what (Caulley 1983). Moreover, this method admits combination with other qualitative methods, most typically with TA (Bowen 2009) which can help identify information inductively. In this thesis, DA uses TA of the interview data and a formal evaluation checklist as tools to analyze the reports and each technique has its particular set of analytical questions based on the suggestions highlighted in TA and DA theory which are next described.

In general, in DA it is recommended to take into consideration that while the written word may be regarded as a superior method to oral data (Caulley 1983), Bowen (2009) warns that documents need not be treated as perfect and complete texts and that both the presence as well as the absence of information may be meaningful. In this regard, Patton (2002) adds that if it is chosen to describe only what was observed, then, “a question might be left in the mind of the reader about whether the other activities had occurred but had simply not been observed” (p. 295).

Another characteristic of DA is that the meaning that unfolds from the analysis of documents must be understood in its own context and with sensitivity (Bowen 2009), so in order to address this issue, in this thesis, all possible data inside as well as outside the document that may contribute to the understanding of the document will be taken into consideration, for example, information in relation to the process of collecting reports.

With regards to the analytical procedures involved in DA, the three main steps suggested by Bowen (2009) are followed. The first (1) involves a glance at the document skimming through the data. In a second turn (2), a more exhaustive reading is done, at which point, an analytical description (3) is started about the broader characteristics of the report, for example, the structure, the organization of information, and the length of the document. Other dimensions in this second phase of the analysis are

(4) the original purpose of the document;

(5) the target audience;

(6) whether the document covers the topic completely or partially and,

(7) whether the document is balanced or uneven in the degree of detail that some aspects of the topic are presented (Bowen 2009).

In this phase, the Evaluation Report Checklist (Miron 2004) from *The Program Evaluation Standards* is used as a reference to comment on these aspects of the report.

The third step in the analysis of the reports involves a deeper evaluation of data, which involves organizing information into categories according to the issues raised in the research questions, supplemented by material from the interview data. In Bowen's own words: "the process involves a careful, more focused re-reading and review of the data. The reviewer takes a closer look at the selected data and performs coding and category construction, based on the data's characteristics, to uncover themes pertinent to a phenomenon" (p. 32).

As far as the seven (7) reports under analysis in this metaevaluation, they are metaevaluated as they were sent to the researcher and treated as case studies (i.e. CS1, CS2, CS3, CS4, CS5, CS6, and CS7). Some decisions had to be made regarding the process of their qualification as actual reports since they presented a diversity of characteristics in terms of content. As an example, the documents received from one specific university could

be considered valid data only to the extent that they demonstrated that the institution had a review policy in place since they were not an actual review report. Therefore, they could not qualify as a case study. As another example, one of the Appendices within one of the review reports received constitutes a report itself and therefore it was qualified as case study (CS3) in order to maximize the use of data.

Besides the DA guidelines described above, the analysis of the reports followed the structure of the Evaluation Report Checklist (ERC) (Miron 2004) since it “draws upon and reflects The Program Evaluation Standards “ (Joint Committee on Standards for Educational Evaluation, 1994). The author of this checklist advises that its purpose is to guide the discussion between the evaluator and the clients about preferred contents of the report and/or to provide formative feedback to report writers and also advises against the use of this checklist as a metaevaluation tool given the variety of needs of clients. In this thesis, the purpose was to use the ERC as a guideline since it provided structure to the analysis of each of the report’s contents (present or absent) and format, and with that to shed light on the character of the review process and to use those findings formatively to the extent possible.

The ERC contains eight content areas, namely, (1) title page; (2) executive summary; (3) table of contents and other sections that preface the report; (4) introduction and background; (5) methodology; (6) results chapter; (7) summary, conclusion and recommendations; and (8) references and appendices. Each one is itemized in different checkpoints from A to E/F with a space for a potentially added item and another one for comments. The checkpoints are rated according to a rubric that asks whether the item has been fully addressed, partially addressed, not addressed at all, or not applicable. In Chapter 3, an explanation of each of the ERC sections is provided before findings are presented in order to clarify what each section deals with in detail.

The order in which the documents were analysed is not based on any particular preference, except in CS2 and CS3 where the latter was analysed right after the former given that, as mentioned, CS3 is an Appendix in CS2 which helped to connect the contents of one with the other. Moreover, even though the researcher was familiar with the contents of the reports from CS5 and CS7 since they had been part of the pilot study conducted prior to this thesis, the analytical strategy used in both projects was different and therefore the DA of these two reports in the current study was approached afresh with a detached outlook.

2.4.3.4. Thematic Analysis (TA)

To analyse interview data or other qualitative data, methodologies such as TA or content analysis may be used. Both of them aim at finding repeated patterns of meaning but while content analysis tends to quantification, TA doesn't, though it can, depending on the purpose of the study. Furthermore, the units of analysis in content analysis are typically words or phrases, whereas TA works with concepts, themes, issues that may be associated with words, phrases or even whole paragraphs. Other methodologies such as conversation analysis or grounded theory are possible alternatives to analyze verbal and written texts respectively, but these have not been considered in the present study because, as Braun and Clark (2006) suggest (and their suggestions applies for an under-researched area such as LPE practice in Australian higher education), it is best to work with a flexible framework which does not limit the study theoretically and allows the researcher to align the goal of the analysis with the main goal of the study.

The analysis of the verbal data and evaluation reports in this metaevaluation does not focus on the linguistic study of specific segments, as may be the case in discourse analysis,

but on actual content and meaning to shed light on the current trends of ALPE processes in university language programs. The aim is to richly describe and interpret content in order to learn and understand what is going on in the data, and to then characterize these processes and in a second phase, use these findings to propose actions for improvement. With regards to the scope of the analysis, in order to address the general purpose of the study, an overall rich description of the themes was pursued. However, as analysis unfolded and themes developed, it was worth paying special attention to a theme or a group of themes which were particularly recurrent or emphasized. While the main research questions of this metaevaluation are grounded in both Language Program Evaluation (LPE) as well as in Program Evaluation (PE) theory, this project is of an exploratory nature and the area of study is under-researched,.

2.4.3.4.1. Coding for TA

Alkin (2011) provides guidelines for conducting qualitative analysis in evaluation and in particular coding of qualitative data which he considers an iterative process, that is, “you look and then you look again” (p. 183). In general, coding is defined as the process of reducing data to manageable “neat coherent chunks” (p. 181) and codes are labels in the form of words or phrases that are assigned to concepts or ideas relevant to the research questions of the study. The sorting of data in this thesis was done via a computer software called NVivo (QSR International 1999) which allowed the researcher to centralize all the research information in one site thus converting the study into an electronic project. This software allows the building of an index of codes and subcodes while carefully reading and re-reading data and the field notes that are taken during reading and re-reading.

The coding process in this thesis followed Alkin’s (2011) guidelines which involve the following steps: (1) respect the “insider perspective” and understand the labels they use to describe how they view “people and human events”; (2) give names to topics that emerge from the field notes; (3) compose a “code book”, that is a list of the labels revealed through the field notes; (4) summarize codes by examining the similarities and differences between and within codes²⁵; (5) find overlap of codes in terms of commonalities which will lead to overarching themes or patterns and ask the role these themes play in the analysis; and (6) include the description of *outliers*, that is, competing explanations and inconsistencies in the data since these are findings too that provide multiple perspectives. Detailed information about the coding process can be found in Appendix 4.

The analysis of the data conforms to the summative purpose of the metaevaluation which seeks to discover the strengths and weaknesses and the character of ALPE processes. But at a higher abstract level, how does the analytical design framework described above fit in with the taxonomy of subjects of inquiry and of inquiry modes from the neorealist evaluation model on which this metaevaluation is based?

2.5. Connection of Analytical Methods with Overall Metaevaluation Framework

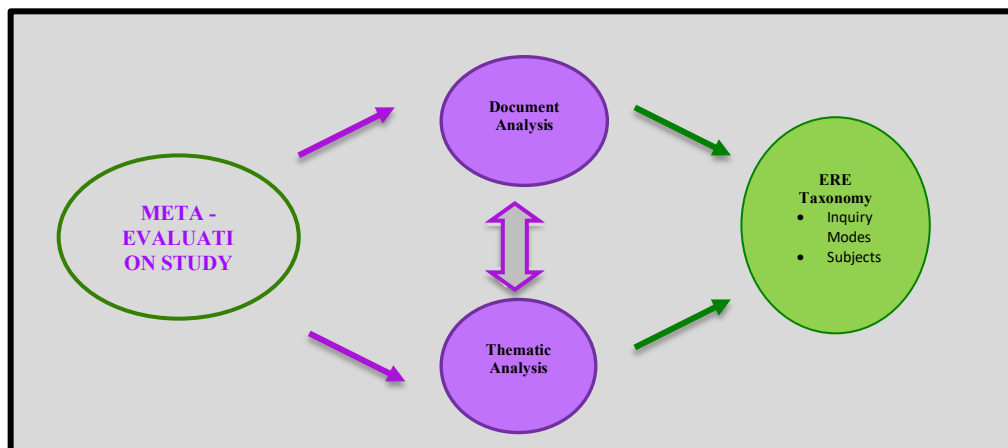


Figure 3. Linking multiple methods

The connection between TA, CS and DA and this metaevaluation's broader framework of the taxonomy of subjects with their categories and subcategories displayed in Figure 3 is established via the analysis of each data item (reports and interview data) which are first described and then cross-examined across all evaluation cases. The subcategories include (see Table 1 below) on the one hand theoretically-driven criteria and also other fundamental issues in evaluation. These criteria are used in the same fashion that Stufflebeam (2001b) suggests in order to gain knowledge about the extent to which the evaluation under study addresses the subtopics that are pointed out in the criteria, thus allowing us to assess the quality of the evaluation.

Furthermore, the process of selection, description and categorization of issues through TA, DA and CSs and the taxonomy of subjects is iterative. First, an attempt was made to learn and understand individual meanings. Second, possible relationships between meanings were established to further classify issues into more abstract themes. Finally, those meanings were reduced into overarching themes. Tables 2 and 3 synthesize the way the thematic process of analysis described above concurs with the taxonomy of inquiry modes, specifically with description, classification and values inquiry.

On the one hand, description aids in detecting issues at a semantic level. For the classification mode, the case studies are examined individually via the observable underlying patterns found through information provided by the description mode. Since in the pilot study, the patterns revealed the important distinction of summative and formative types of evaluations, in this thesis they are used as the two underlying categories in the classification mode. Hence, the information obtained from the review cases through the description mode will be classified according to the attributes of each type of evaluation.

	Evaluation context	Evaluation activities	Evaluation consequences	Domain Specific
Concept	Circumstances (organizational and programmatic) within which evaluation occurs	Procedures in planning, carrying out, and disseminating evaluation	Changes that do (or do not) occur as a result of evaluation	Questions within a specific area of evaluation practice
Subcategories	<ul style="list-style-type: none"> • Commission • Purpose • Evaluators • Policy 	<ul style="list-style-type: none"> • Approaches • Components • Practices 	<ul style="list-style-type: none"> • Use/influence • Participants • Evaluation context, context activities, and professional issues 	<ul style="list-style-type: none"> • Essential issues in successful evaluation practice in LPE settings

Table 2. Taxonomy of subjects for a metaevaluation of ALPE

On the other hand, with regards to the study of values, in the integrative framework by Mark, Henry and Julnes (1999), sometimes values from certain stakeholder groups are more predominant than others. Hence, in this metaevaluation the analysis involves the systematic study of values by finding out first who the stakeholders are; second, the nature of those values; and third, whether the values of a particular group are prioritized in the evaluations under study. Values inquiry thus may assist in the reasoning about the findings from the description and classification modes.

DESCRIPTION (directly detectable)	CLASSIFICATION (underlying mechanisms)	VALUES INQUIRY (complexity and stratification of reality)
TA's semantic level	TA's latent level	TA's latent/semantic level

Table 3. Inquiry Modes and Thematic Analysis

After each evaluation case has been analysed, the metaevaluation then establishes a comparison across the evaluations in order to synthesize the findings which, in turn, helps gain an empirically-based “picture” of the current character of evaluation practices in Australian university language departments.

Given the amount of data, both documentary and interview, generated in this thesis, once the analysis of both sources had been completed, a decision was made to prioritize reporting the material from the Document Analysis as it gave a more comprehensive picture of each case study. For this reason, the full results of the Thematic Analysis are not presented in the thesis, given the constraints of length. Instead, the role of the interview data was to supplement, illustrate and contextualize themes emerging from the Document Analysis.

2.6. The Role of Context in Metaevaluation and Implications for this Thesis

Context represents one of the subject categories to be studied in Mark's framework and special attention is paid in this metaevaluation to the policy/organizational context in which the reviews under study are embedded. Moreover, the contextualization of findings is essential to provide credible and valid interpretations, as discussed in 1.2. As public educational institutions are expected to be accountable for how public money is spent, accountability seems to be the driver of QA processes today, particularly focusing on how the quality of education is managed and improved. However, the concept of 'quality' becomes subverted and the character of reviews in the current NPM model of Australian universities runs the risk of being void of any evaluation theoretical foundation. In light of the accountability-free power structures of the NPM ideology managing universities, metaevaluation becomes an imperative process, not only to ensure findings are the result of careful evaluation studies, but also as a checks and balances tool, so that issues of an ethical and fairness nature are the centre of academic and institutional governance discussions as needed.

Chapter 3: Findings

Chapter 3: Findings

This chapter is a detailed explanation of our metaevaluation findings present in different sections deriving from the Evaluation Report Checklist (Miron 2004) (see 2.4.3.3 above). These include specific key context (i.e. aligning languages Schools and programs with University priorities) and supporting evidence (Tables of Contents (ToCs), etc.), in addition to findings about our research on the evaluand and evaluation context, methodology, results, summary, conclusions and recommendations. Throughout, material from the interview data is integrated with the results of the analysis, in order to contextualize and supplement the findings. The aim of this presentation of the findings of our Case Studies is to identify evidence in support of the discussion in Chapter 4, which focuses on Research Questions 1 and 2, and in Chapter 5, which discusses Research Question 3.

3.1 Executive Summary of Reports' Findings

Before using the ERC to guide us through the findings of the reports, and although the ERC does not deal separately with the Executive Summary (ES), in this section a description of the ES in the reports is presented. The ES is a feature of the majority of the reports. CS3 and CS7 lack an ES whereas in the rest of the documents studied, the ES presents uniformly a summary of findings, a brief description of methodologies, and recommendations except for CS4, where no methodologies are summarized. It must be noticed that in the ES in CS2, CS4 and CS5, the third person is used to transmit what the Review Panel (RP) "considered", "examined", "commended" and "recommended".

The implications of findings are not made explicit but the general emphasis on recommendations may imply a certain pressure to implement the action/change suggested as shown in the following quotes:

While some discontinuations are unavoidable, it is important to lower the overall rate of discontinuation and increase enrollments in the DML. (CS1)

Whilst the report lists a broad range of commendations, aspects worthy of particular note include: [...], Similarly, the report outlines a range of recommendations for consideration by SLC and the Faculty. (CS2)

Importantly, the review yielded some very favourable findings and culminated in a series of 13 recommendations that are designed to strengthen the Program, ensure its continued success, and identify areas of potential growth.

The Diploma in Languages faces a set of unique challenges. [...] Indeed, there appear to be some very realistic opportunities for growth [...] Similarly, there are numerous opportunities for improving [...] Overall, while there are some opportunities for development and improvement [...] (CS4)

In order to raise the education and research profile of the school with a view to taking its rightful place, the panel makes some significant recommendations in these areas. (CS5)

We urge the consideration of these recommendations, and the detailed analysis which follows.²⁶ (CS6)

The summary of findings in every case usually starts with positive findings followed by negative ones, and generally negative findings are conceptualized as challenges or opportunities for improvement.²⁷ As a distinguished feature, the positive engagement of program staff with the review by which the RP “was impressed” is presented as a positive finding in CS5. In sum, the language and content of the findings, as with other areas of the report, are particularly conditioned as we shall see by the managerial context affecting the overall purpose of the review exercise, with a relative backgrounding of educative perspectives and values.

3.1.1 Program Description

A program description is absent from the CS2 and CS5 ESs and in the rest of ESs, program descriptions vary in length. For example, in CS1 a very brief description is found in

comparison to those in CS4 and CS6. CS6's ES presents the most detailed descriptive paragraph about the program²⁸ which is further complemented at the end of the section highlighting first of all the relevance of the program not only within the educational framework of the university but also within national strategic plans, thus denoting a rather justificatory tone.

3.1.2 Brief description of methodologies and analytic strategy

CS3 does not provide any methodological details anywhere in the report and in CS7 they are mentioned in the body of the report (the results chapter). As for the reports with an ES, in the ES of CS4, as mentioned, does not provide any methodological details, and while the methodological references in the rest of the ESs (that is, in CS1, CS2, CS5 and CS6) vary in the level of detail (e.g. data sources and data collection methods such as documentation, submissions, interviews, meetings), generally, there seems to be an emphasis on range and quantity of stakeholders participating in the review as informants, with the exception of CS6. For example, CS1's description focuses on informants, providing the exact number for each set (e.g. students and staff) participating in interviews. In CS2, the description includes data sources (e.g. School's self-review and additional documentation),²⁹ the time spent in meetings between staff and the Review Panel, and an emphasis on quantity and variety of stakeholders (e.g. "numerous stakeholders"). In CS5, as in CS2, emphasis is placed on stakeholders (individual, groups or organizations) whose perspectives are collected via submissions, meetings or information sessions. Differently, in CS6, a brief description places the stress on data sources used (e.g. "income sources and modelling costings") and data obtained through benchmarking "internally and externally".

3.1.3 Evaluation questions and evaluation purpose(s)

Whether explicit or implicit, CS2, CS4, and CS5 are the only reports with evaluation questions and purposes in the ES whereas in CS1 and CS6 these are introduced in the Introduction chapter which is another chapter where these checkpoints can be included according to the ERC. In CS3 and CS7, neither questions nor evaluation purposes are made explicit anywhere in the reports, and therefore, depend on inference.

3.1.4. Areas of study

In CS1, CS4, and CS5, the ES mentions areas about which questions may be asked which denote the use of these as a framework to guide data analysis. For example, in CS1 a question is asked about the area of student discontinuation, which is expanded in the Results chapter into a wider framework of questions. In CS4 and CS5, certain areas that the panel considered about the evaluand are mentioned. For example, in CS4 the writer mentions “among other things” enrolment figures, structure and content of the evaluand, its target audience and marketing, teaching methodologies, student and community satisfaction and student outcomes. Similarly, in CS5, the areas that are mentioned are internal organization, governance, resources, staffing, staff development, linkages with other areas of the University and outreach to the broader community³⁰. While, in CS2, no specific areas are mentioned, “the overall performance of the School” may refer to the areas in CS2’s ToR since these correspond to the ones in the sections within the Results chapter. A similar pattern is found in the other two reports without an ES whereby CS3 is guided by a set of areas which are identical to those in CS2’s Terms of Reference (ToR)³¹,

and in CS7, the topics in the headings and subheadings in the Results chapter could be considered policy mandated areas guiding the analysis, too.

An analysis of all the areas/questions identified reveal certain similarities. Apart from the above mentioned areas in CS2 and CS3 bearing a strong similarity, three broad areas from CS2 and CS3, namely, governance structure, research performance and outreach to the community, are shared also by CS5's framework, and the third one in particular, appears to be an area of interest in CS7 too ("External Stakeholder satisfaction").

Regarding the rest of the case studies, some shared areas can be identified. For example, in CS1, CS6, and CS7 the focus seems to be on student numbers, and while in CS7 "viability" is one of the areas to be studied, in CS6, the focus of the questions on retention, attrition rates, enrolments, teaching costs and deficit are clearly associated with finances. Therefore, it can be argued that "retention" in CS7 and "discontinuations" in CS1 may also be areas associated with financial viability. In CS5, student numbers are also inquired about two particular areas, "taught Masters programs" as well as students studying a "Higher degree by research", therefore the "Masters coursework programs" in CS2 may also be of interest from the point of view of student numbers³².

A shared area among the case studies except in CS6 is the academic aspect of the evaluand. However, this apparently educative perspective is better read as an aspect of managerial concerns. For example, in CS1, the question as to "why" students discontinue denotes an interest in finding out possible weaknesses that may lead to discontinuations. In CS7, similarly, the "Academic design" of the program is also an area of interest to be studied. Interestingly, in CS4, while the focus of all its areas is clearly "language teaching", an obvious relationship is established between student satisfaction and "demand and growth" opportunities, and therefore CS7's three areas on satisfaction at three levels (e.g.

“current student”, “graduate” and “external stakeholder”) may also be connected with increasing student numbers. Moreover, since graduate outcomes including “destinations and employment” in CS7 are in turn connected with graduate satisfaction, these outcomes may be important to be considered in order to retain or attract students, which is an area that emerges in CS4 (“Destinations and outcomes for graduates”). In CS2 and CS3 the “first year experience” (the initial encounter with the undergraduate program) is stressed.

The area of “diversity and equity” is found in CS2 and appears also as an area to be studied in relation to the viability of the program in CS7.

In CS2 and CS3’s “joint teaching” seems to be an area of interest also in CS4 (“cross-institutional”). While there is no evidence in the denomination of areas that can lead to relating this topic to student numbers, it must be noted that since all the aspects in CS2 and CS3’s “Management of undergraduate and postgraduate curricula and teaching”, namely, “first year experience, diversity and equity, masters coursework programs and joint teaching”, it would be unlikely to assume that “joint teaching” is the only aspect reviewed under this section unrelated to student numbers.

There are some areas of interest in at least two of the reports namely, CS5 and CS7, which appear to be primarily educative: on the one hand “collaboration”, or “linkages” under CS5’s “Disciplinary Balance” within the School and “other areas of College and University”, and on the other, the “teaching-research nexus” which appears in CS7 and also under CS5’s research performance (“integration with educational programs”).

In general, however, from the description of the ESs presented here it appears that while the names of areas and aspects within areas denote an interest in academic matters, overall these are studied from an administrative perspective, and particularly, the focus on student numbers associated with viability stands out in some of the explicit names of the

areas (e.g. discontinuations, growth and demand, increase of student numbers, increase enrolments), as well as in the implicit relationship between student numbers with areas such as student satisfaction, graduate outcomes and employment, diversity, equity, retention, etc.

In spite of the different location in the reports where these areas together with evaluation questions and purposes are presented, this important checkpoint is next fully tackled in order to identify possible similar patterns of meaning.

3.2. Aligning the languages Schools and programs with University priorities. Context in CS2, CS4 and CS5.

In the following three reports, CS2, CS4 and CS5, the evaluation questions identified in the ES are indirect and closely related to the job expected of the reviewers from which a purpose for the review can be inferred. This evidence suggests a broader underlying purpose to use the review in these three case studies to align the languages Schools and programs with University priorities whereby the “context”, that is, the Strategic Plans, serve as the criteria to measure the extent to which the evaluand is aligned with such priorities. The ideology of New Public Management underlying the evaluations is particularly clear in these cases.

In CS2, the writer states that what “the panel considered” was “the overall performance of the School as well as the various opportunities and challenges”. Similarly, in CS4, the areas of the program that “The panel examined” and the recommendations “designed to strengthen the Program, ensure its continued success, and identify areas of potential growth” denote two broad questions. In the ES of CS5, this same pattern can be identified where the writer describes what the Review Panel “considered”, referring to the

areas mentioned above, and its “developmental recommendations” “to raise the educational and research profile of the School” “to aspire to national leadership”.

Therefore, from this evidence on questions, a purpose in CS2, CS4, and CS5 is suggested for the improvement of the areas selected through the implementation of the RP’s recommendations, and in CS5 particularly, to make the School more competitive.

However, an additional purpose can be identified in the ES of CS5 which is “to set future directions in both Education and Research”. But what are these directions based on? The ES does not offer an explanation but evidence in the ToR connects these directions with a “context” which, as a matter of fact, CS2 and CS4’s ToR present too, as one to be taken into account by the reviewers when reviewing the areas in the analytical framework described above. This “context” is associated with the “Strategic Plans” of the universities, though no explanation or description is provided other than a specific date for their implementation and the different levels of the plans. For example, in CS2, the breaking down of these plans shows three levels, namely, University, Faculty and School, and in CS4, similarly, a “Top Level Organisational Structure” is distinguished from “Strategic Plan 2008 – 2012”, and from “Faculty of Humanities and Social Sciences Business Plan 2009-2011”³³. In CS5, while the “context” to be taken into consideration is explained very broadly (the context of a College of Arts and its linkages with another College, and the context of “the wider framework of the disciplines of Languages and Linguistics in the Australian university sector”), the provision of a framework of evaluation questions under the areas selected to be reviewed allows us to discover that not only do concrete “College and University” Strategic Plans exist but also that they are to be used in order to articulate certain areas (“taught graduate programs”; “research performance”; and range and delivery of courses) with them.

3.2.1. Evaluation questions and evaluation purpose in CS1, CS6, CS3, and CS7

The questions found in CS1 are, as in CS2, CS4 and CS5, derived from the task assigned to the reviewers (e.g. “we were tasked to do”), namely, “to provide a snapshot of the state” of the program “including its strengths and weakness³⁴ and opportunities for improvement”, and additionally, derived from what “the primary concern of the review was” (“discontinuations”) to also, as commented above, “understand the reasons why students might discontinue”. Interestingly, while these indirect questions imply using the review for knowledge production, that is, to gain a deeper understanding about what causes student discontinuation, in order to improve the program, CS1’s ES adds the goal of the recommendations to “lower discontinuations *and* increase enrollments” .

In CS6, the self-reviewers suggest that the analysis has been guided by the questions in a “Review Notice” (“We have responded in detail to the questions and allegations in the review notice, documented current plans and developed new plans”³⁵). Interestingly, two contending lines of information may be inferred from CS6’s Introduction chapter where the evaluation questions and evaluation purpose appear. The first one can be proved through the existence of the Review Notice (RN) mentioned in the ES, which “was notified formally” to program staff, with triggers, Terms of Reference and evaluation questions. The second one, in contrast, gets revealed through CS6’s self-reviewers’ own sets of “aims”.

The “terms of reference” of CS6 reveal two tasks³⁶, namely, “To review the viability” of the program and “to make recommendations” from which a purpose to improve the viability of the program through suggested actions in the recommendations may be inferred. However, additional contextual information in this section points out an implicit use of the review which confirms the complex initiation of the review.

3.2.2. Recommendations

Recommendations in the reports with an ES (CS1, CS2, CS4, CS5, CS6) appear either summarized or in full. For example, in CS1, the most “urgent” are distinguished from “other significant” improvements, and in CS2, not only are recommendations of “particular significance” included (8 recommendations) but these are preceded by “commendations of particular note” (4 commendations) selected from “a broad list”. In CS4, similarly, two commendations precede recommendations but, in this case, the full list is provided (13 in number) and separately. In CS5, all the recommendations (18 in total) are provided in a list as they are in CS6 too (2 recommendations). Furthermore, as noted above the purpose of the recommendations is indicated with the exception of CS6 where the only comment provided is that both of the recommendations are “strongly recommended”³⁷.

It must be noted that the independent lists in CS4 and CS5 do not offer any explanation about the recommendations (though in CS5’s ES these are qualified as “developmental”) nor are they structured in any particular way apart from the fact that the numbering and the order in which they are listed matches the numbering and the order in which they appear at the end of every section in the Results chapter. However, in CS5’s list, particularly, some irregularities can be identified. Hence, while the list reflects exact names used for sections and subsections in the Results chapter, which in turn reflect the specific topics asked in the evaluation questions, this structure is gradually abandoned. Moreover, subdivisions are numbered only in the first category for no apparent reason as Table 4 shows.

1. Organizational and Governance Structures	1.A Organization and governance
---	---------------------------------

	<p>Recommendation 1.1</p> <p><i>1.B Administrative support</i></p> <p>Recommendation 1.2</p> <p>Recommendation 1.3</p> <p><i>1.C Facilities and staffing</i></p> <p>Recommendation 1.4</p> <p>Recommendation 1.5</p> <p>Recommendation 1.6</p>
2. Disciplinary balance	<p>Recommendation 2.1</p> <p><i>Linkages within the College</i></p> <p>Recommendation 2.2</p> <p><i>Linkages with College of Asia and the Pacific</i></p>
3. Range and delivery	<p>Recommendation 3.1</p> <p>Recommendation 3.2</p> <p>Recommendation 3.3</p> <p><i>Honours</i></p> <p>Recommendation 3.4</p>
4. Capacity for further development of taught Masters programs	<p>Recommendation 4.1</p>
5. Higher degree by research	<p>Recommendation 5.1</p>
6. National and international research performance	<p>Recommendation 6.1</p> <p>Recommendation 6.2</p>
7. Outreach to the community	<p>Recommendation 7.1</p> <p>Recommendation 7.2</p>

Table 4 Structure of Report and Recommendations, CS5

In CS3 and CS7 which lack an ES, recommendations are presented in the Results chapter and in CS7 also at the end of the report before the Appendices.

3.3. The Table of Contents (ToC) and Other Sections that Preface the Report

3.3.1 Description

The Table of Contents (ToC) provides a quick view of the contents inside the report and facilitates the finding of information and the understanding of the variety and quantity of material inside the report. Thus, according to the Evaluation Report Checklist (ERC), the ToC of an evaluation should have “at least” first and second headers, accurate page numbers and titles, and, “if appropriate”, a list of tables and figures, acronyms and abbreviations, and appendices. Finally, an additional section introduced in this section is “acknowledgements” referencing all of those who have contributed both to the evaluation process such as “sponsors, data collectors, informants, research assistants, reviewers of the report, etc.” as well as “contributors to the report”.

3.3.2. Findings

In this study, CS2 and CS7 lack a ToC and therefore It was necessary to infer what a ToC would have looked like for these reports. Moreover, of the five reports with a ToC, only CS1, CS4 and CS6 have at least first and second header levels, and CS3 seems to prioritize Appendices since they are the only part in the ToC with first and second header levels. Regarding titles and page numbers, in CS1, CS5 and CS6 these are all correct whereas in CS3 there are no page numbers for main contents or appendices, and in CS4, page numbers are inaccurate in three instances.

3.3.3. List of Tables, Graphs and/or Figures; Acronyms and Abbreviations;

Acknowledgments

C1, CS6, and CS7 are the only reports where tables, graphs and/or figures are used but CS6 is the only report with a separate list of these also indicating page numbers which are all accurate.

Furthermore, no report presents a list of acronyms and abbreviations though all reports use to a higher or lesser extent acronyms and abbreviations. Moreover, the full words are not indicated in any of the reports except in CS3 (e.g. “The Postgraduate Research Support Scheme (PRSS)”) which required research in order to learn what the initials stand for.

With regards to acknowledgments, the only reports where acknowledgments are included are CS2, CS4 and CS5, though they appear in different sections. In CS4, for instance, acknowledgments are conceptualized as a thank you note and it is presented at the end of the “Review Process” section where the “assistance in the successful conduct of this review” by the staff and management and external stakeholders and students who participated as informants in this review is acknowledged. In CS2, more than acknowledge, the anonymous writer praises the good predisposition of staff to engage with the review which in fact is presented not once but three times in three different sections, namely, the ES, the “Introduction” and “Summary/Conclusions, as evidenced in the following quotes:

The panel was very impressed with the positive and constructive way in which most members [...] engaged in the review process and the openness with which issues were discussed.

The panel was very impressed with the positive and constructive way in which most members of SLC engaged in the review process and the openness with which opportunities and challenges were discussed. Of particular note was the strong collegiality that was evidenced during the review.

Finally, the panel wishes to acknowledge the constructive engagement of the School in carrying out the review and thank those who contributed so positively and openly to the review.

As mentioned, this topic also emerges in CS5's ES where the participation of staff in the review is framed within other positive findings³⁸.

3.3.4. Additional Section: The Terms of Reference (ToR)

Apart from the Terms of Reference (ToR) in CS6 which are explained by the authors of the report in the Introduction chapter, in CS2, CS4, and CS5 a whole independent section called "Terms of Reference" is included which contains also the composition of the Review Panel.

But what exactly are Terms of Reference? What is the function of a ToR in an evaluation report? Do they offer the same kind of information across the reports? And also, why do some reports come with ToR while others such as CS1 and CS7 don't?

Since there is no chapter in the ERC under such a term and no evidence has been found in the Evaluation literature about ToR, an answer was sought by contacting a renowned Evaluation researcher (Robert Stake) who through an email message exchange admits not to have any knowledge about ToR ("I have heard people say "terms of reference" and I too didn't know what it meant") and resorts to an online search (Wikipedia) to find out a definition which he finds "sensible" with "important" elements such as

- vision, objectives, scope and deliverables (i.e. what has to be achieved)
- stakeholders, roles and responsibilities (i.e. who will take part in it)
- resource, financial and quality plans (i.e. how it will be achieved)
- work breakdown structure and schedule (i.e. when it will be achieved)

which he goes on to explain is used by "bureaucracies and administrations".

On the whole, the focus of the ToR in CS2, CS4, and CS5 is mainly on instructions for the reviewers as reflected in the following expressions: "Having considered [...], the review

panel should [...]” (CS2) or “the Panel will take into account [...]” (CS4); “Specifically, the Review Panel is to undertake [...]” (CS5). However, the range of instructions and the level of depth and detail is far from what the online definition suggests. Hence, apart from the “context” explained above revealing Strategic Plans (which in CS5 is suggested in the ES instead of the ToR), instructions about data sources and collection methods are provided which generally tend to be brief and at times quite vague. For example, in CS2 only one data source and the areas to be examined are mentioned; in CS4, the name of stakeholder groups whose views are to be taken into consideration, the areas to be examined and a specific restriction form the core content of the instructions; and similarly, in CS5, reviewers are asked to review certain areas, to “consult widely”, as well as being told who to report to.

Furthermore, the areas to be studied are presented with more or less detail as the following samples show:

- Evaluate the quality of the School's management of curricula and teaching at undergraduate and postgraduate levels with some attention to issues of first year experience, diversity and equity, Masters coursework programs and joint teaching of programs across schools; (CS2)
- Structure, content, quality and overall coherence of language teaching (CS4)

whereas in CS5, specific questions (from 3 to 5 in number), follow the denomination of areas. Similarly, according to CS6’s Introduction, the ToR reflect the main area to be reviewed (“viability [...] with respect to enrolment numbers and financial outcomes”) also followed by three specific questions.

Data sources and collection methods in the ToR do not always reflect what is found in later sections. However, while CS6’s authors reflect this as a conscious decision

We will address each of these questions [Review Notice’s evaluation questions] separately, with extra data and information in Appendices

in CS2 the only methodological reference in the ToR which reveals a unique data source for the panel to consider (“the School’s Self-Review documentation”) is expanded in the

Introduction

The panel was provided with a detailed self-review undertaken by SLC and relevant additional documentation [...] Submissions were also called for [...] Over the space of three days [...] the panel met with numerous stakeholders [...]

without any justification, thus suggesting, either slack practices and/or a conception of the report for strictly internal use.

The one topic that seems to be emphasized in the ToR of CS2, CS4, and CS5, which was also found in the ESs, is the variety of stakeholders to be consulted, which in CS2 and CS5 is brought up in other sections as well.

An additional metaevaluation finding from the ToR are triggers, that is, things which initiate the review process, which might seem to indicate an educative focus of quality assurance. Apart from CS6’s particular triggers which lead to key metaevaluation findings revealing a political use of the review, periodical cycles of reviews³⁹ appear to be the trigger in other reports though this is not always reflected in the ToR. For example, in CS4’s ToR “the University’s five-year program review cycle” is mentioned, while in CS2, the writer states that this review is part of an “inaugural five-yearly School review” in the ES. Similarly, in CS5, the writer states that the review is “part of an ongoing program of quality assurance in Education” also in the ES. CS7 equally has a cyclical character though this is only learned through the interview data.

However, thanks to verbal data from interviewee (I) who participated in CS5 as a member of the staff in the School (“I was involved”), evidence is found proving a much more complex trigger that implies underlying predetermined evaluation purposes. Hence, while

(I) confirms that “this is embedded in the rules” and that these processes are done on a rotational basis and vary in terms of evaluands:

14. [...] There are, I mean, periodic reviews of, of schools or, or er MOUs⁴⁰, administrative entities within the university [...]

he also points out the targeted character of reviews since these are used as an institutional push for changes, that is, by managers who have been specifically appointed to target particular strategic objectives:

12. [...] In addition, here at (U14) er since— abo- about 4, 5 years ago there was a new er deputy vice chancellor or, well, a deputy dean education who er came with an agenda of change that also asked the language programs to be benchmarked, an international benchmark, and so I was the convener of the (L) program so I was also involved in this.

In fact, according to (I), there is an anticipation by staff about imposition of changes ensuing from reviews in general and from CS5 in particular which, in turn, are associated with an erosion of programs’ teaching capabilities through the reduction of resources, thus affecting their educational quality:

16. [...] an imperialistic move to change the er practices that were maintaining— practices that were maintaining the quality of language teaching against practices that were more— perceived more as more commercial or more er towards er sort of aligning our university with practices at other universities, including Group of 8 universities, but which people perceive them as not, not good.

115. [...] I mean, it was, at some level, the imposition of change and the review was, was rightly perceived by staff as a, as, as an instrument that was going to be used to impose change that was not welcome.

1190. [...] the (X) programs were reviewed, because it was essentially to impose a lot of cuts and changes.

Furthermore, while (I) associates reviews generally with making programs become “restricted and casual”, in CS5, the imposition of changes was also associated with additional institutional priorities regarding research:

160. [...] the review also was geared very much to align our school with the needs to be very successful in the incoming ERA, the Excellence for Research in Australia

which required staff to

160. [...] produce research as a priority and not to put so much time into teaching.

In turn, the need to push goals to produce research seems to be in line with other university measures connected with an overarching university strategy that aims to secure ranking positions and for which another designated 'hire' was brought in:

160 [...] But from the top, I mean, there was also a change of vice chancellors and the change of vice chancellor meant that we have gone from having someone like (P23), who was a very experienced player in the field, to have someone like (P24), who is someone who was able to bring er one university, that was (U17), from nothing into the— all the research charts that are used [and so basically—

While understanding the need to do research, for (I), reviews, including CS5, are used to push “overarching issues” such as ““You had to comply with the ERA [Excellence in Research in Australia, similar to the British REF], you have to be prepared for this, you have to do that and that and that”” that force programs to improve but in conditions that are not conducive to good practice:

1132. [...] you are forced to work er under constraints that are completely unsuitable for the language teaching environment

and which have to be done at *any* expense which admittedly includes subversive practices as the following quote suggests:

166. Learn to live with and, and it's a very difficult thing because, I mean, at a lot of levels er what we have experience is, of course, I mean, 'reduce number of tutors er and, and er you have to cut corners here and there, what is the least worst outcome in a bad outcome' and all this kind of things, but it's, I mean, they are trying to fulfill the—

Q67. They have an agenda, is that right?

168. Well, of course, they have their agenda, they have their agenda of, of having this, and this, and this, and this parameter to be improved and it doesn't matter how it improves, it has to be improved. And so er at the level of the language teaching, we have been able to maintain a lot of the quality but using a lot of ingenuity and a lot of imagination, simply because, I mean, the restrictions are big.

In conclusion, (I) summarizes that in this context, reviews, and CS5 particularly, was a “push to impose” within a broader scheme of changes which has reduced the program considerably and which leads to cynicism:

I31. [...] I mean, this is my personal view but er the review was used as a tool to change things and as a consequence of the review, this— I mean, and, and together with the review, there was an overarching change of the curriculum in all the university but specially in the faculty of arts that was the biggest faculty in the, in the university, and there, we lost the position of having in (D1) a compulsory major. And so this— I mean, today the only program that has a compulsory major in the language at (U14) is the bachelor of languages and I’m the (PC1) of it and I usually joke that er I still have to fight with administrators that will, some day, come to me and say that ‘I don’t have to have a compulsory language major in the bachelor of language’.
(laugh))

However, while the pressures to reduce costs are considered “top down pressures” (“You have to reduce the cost of this, you have to reduce this, you have to reduce the number of tutors, you have to reduce—’, and so on”), in CS5, a personal factor is suggested since these were demanded by a prejudiced manager with a particular ideological slant about how funding should be allocated or earned:

I116. [...] And so this is, this is a constant battle in which you have to resocialize the people on top on the need of language programs constantly, when you— when it happens that you have someone like we have, that have a lot of prejudice, didn’t— was not permeable to socialization and have a lot of constraint, then what you get is a push to impose, a colonial attitude, essentially it’s a tool of colonialism of the other disciplines that result on a lot of pain because essentially language teachers will not relinquish the issue of teaching the language well, but they will have a lot of extra impositions that they will have to comply with.

I118. Yes, but I mean, but the problem is that it depends on who you are talking to, because, I mean, we have— I mean, if we had had the same review under the previous dean that we had, who was a (N7), bilingual (N7), who knew that you have to have a lot of effort to get— so if someone it’s already— has a grasp of the issues, then it’s not as difficult, because you don’t have to sort of—

I146. [...] I mean, there are people who are knowledgeable, there are people who understand but you can have the bad luck of having someone who doesn’t understand and try to push things, I mean, I’ll give you just a simple example: Diploma of languages. Has to be promoted part time because it’s sequential, students cannot,

well, one of the administrators 'so why don't you allow them to do it in one year?' And, it's a very basic issue. You cannot do (L1) 1, (L1) 2 and (L1) 3 in the same semester. It's impossible.

However, to complicate matters further, the targeted nature of CS5 also occurred at a different level since according to (I) the commissioning of the review was taken advantage of due to a "battle", a "clash" between the commissioner, who lacks an understanding of language teaching, and the HoS "who is very militant in maintaining the standards", in order to make a personnel change to facilitate predetermined changes:

I11. [...] parachuted on us, in the sense that it was something that- it was something to be expected but, in reality, it was as a consequence of this era of a long and very difficult battle with the new era sub-dean education who didn't understand the requirements of language teaching.

I13. [...] the clash of these 2 personalities, at some level, resulted on using the institutional possibility of the review to change things.

Interestingly, not only is the review used as a pretext to replace a particular individual resisting changes but the replacement was in fact a member of the Review Panel:

I13. [...] And not only that, they also brought someone from outside as a professor to put it as the new head of school in the next year but the review was conducted before this person took the place but she was part of the committee, she was part of the review.

Q14. Of the panel.

I15. Of the panel [...]

All in all, according to (I), the use of a review to impose causes "a lot of levels of conflict" such as "just having a review that was perceived as an imposition, plus the level within the school on deciding strategies to face this". Interestingly, when a change perceived by staff as needed, such as the upgrade for "teaching with technology", was voiced to the Review Panel, conflict also arose. According to (I), budgetary discipline (e.g. "didn't want to invest on this") seemed to have been the cause of the delay in such a "key issue" since commissioners "were feeling ashamed that you were telling people of the

outside members of the panel how bad things were because of their diligence". Thus, while the outcome was positive ("we won that"⁴¹), it was perceived as a negative experience since "it required a lot of militancy in the sense that, I mean, you have to show them how things were bad, and go against the dean" with even adverse consequences such as a growing animosity between managers and staff ("And so again this produced a rage of (education), but [why you, you—?")

Moreover, in interview data from M, linked to CS5⁴², an additional trigger can be identified which further confirms the targeted character of the review.

M8. [...] the school and the faculty had a reason to be worried about, about the languages [...]

which was known not only by the commissioners but also by the reviewers themselves:

M10. [...] anyhow, there was clearly a problem and we all knew there was a problem and we all knew it in different ways.

But what was the nature of this problem? According to (M) ("this is my private view"), languages at this University "has always been very weak, [...]"⁴³. This weakness gets revealed through a finding ensuing from the review which is making students write their honors thesis in English rather than in the subject language:

M24. [...] Now, that being beside this other fact that they absolutely insist that all the students doing honours write their honour's thesis in (L2). This used to be the case, this is the (L2) group er but it was the same all around. This is particularly the (L2) group.

which in (M)'s view was possible to do in the past whereas in the current system it has little applicability since students are not prepared to pursue post-graduate studies:

M26. [...] Er these days we have people who do precisely 100 credits before they get into honours, some of them starting from beginners, and some of them doing combined degrees. Why write in L2? No good to them: they're not going to be L2 academics, [...].

Interestingly, the writing of the honours thesis in the L2 is tied to another underlying problem, that is, the reduction of the number of students doing their honours year (4th undergraduate year in Australia)

M26. [...] and everybody thought this was the problem, that if they're not getting enough honours students is because of this. They're just losing their honours students.

which (M) perceived was the purpose, or at least one of the purposes, of the review.

M12. [...] to increase um their high degrees and their honours students and make more sense of what they were trying to do.

Therefore, it seems that rather than an intellectual problem, the perception of the underlying purpose of increasing high degree numbers (honours) for the review may in fact be the cause of the program's 'insistence' since it may imply, as commented by (I) a demeaning of the quality of the program. Interestingly, the argument of disciplinary weakness among reviewers seems to be a perception that commissioners take advantage of since (I) argues that a completely different argument was used to press for this change:

I84. And so, when we discuss this issue, for instance, it was a very difficult issue to discuss because the things that were coming back to us was that other universities were doing it.

which denotes alignment with unknown broader goals and the need to compete for students with other universities.

In conclusion, these metaevaluation findings show that behind the cyclical quality assurance trigger indicated in CS5's report there lies the political use of the review including (at least in I's view) the self-interested selection of reviewers to align the evaluand with specific University priorities which are associated with strategic goals such as research outputs and quality of language teaching while applying budgetary discipline which compromises the quality of language teaching, and making the program more easily

accessible to make program competitive in order to attract students, thus increasing enrolments in line with NPM principles of basing courses on commercial viability.

Finally, in CS1, though the report does not contain Terms of Reference nor any evidence suggesting cycles of reviews, the “Brief and Scope” section states the existence of a previous review from which the current one is “a follow-up study”. This thus suggests that the results from the first review may be the primary trigger for the current review.

Since no explanation is found in CS1 as to why there was a need to conduct a second evaluation using qualitative data in contrast with the previous quantitative oriented review, a potential reason could be a negative reception of findings from the quantitative report which may have prevented any commitment to actions (known in the evaluation field as evaluation utilization). While only further research could confirm this and also shed light on what may have caused such an opposition, the commissioning of this consecutive review suggests that the commissioners may have perceived the quantitative orientation of the first review as a potential trigger of the negative reaction, thus commissioning one based on qualitative data. Moreover, two further evaluation issues were reconsidered and changed from the first to the second review as information in the “Brief and Scope” section indicates. The first one concerned the reviewers. Thus, while the identity of the “preparer” of the first report reveals an individual not affiliated to languages⁴⁴, the membership of the reviewers in the present review reveals an area of expertise pertinent to the languages discipline (e.g. Applied Linguistics). And the second one concerned the source of the commission which changed from “the Faculty’s Academic Programs Office” in the first review to a closer-to-home commissioner, that is, the Head of School. Furthermore, the dates from both reports showing that there was a three-year lapse between the two reviews may too be an indication that perhaps no changes were implemented after the quantitative report and

that the issue of discontinuation persisted. Interestingly, available anecdotal data suggests that the new methodological approach, the change of reviewers and of the commissioning source of the review were not issues affecting evaluation utilization since the findings from CS1 were not well taken either (“they hated it”, according to one interviewee from CS1)⁴⁵.

This evidence, therefore, leads to the claim that the commissioning of the current review with the corresponding modifications may in fact have been triggered by the need to enhance evaluation utilization, that is, as an opportunity to gather evidence of a qualitative nature in an attempt to increase staff’s commitment to address the issue of discontinuations.

Apart from the review purposes explained earlier, a brief description of the evaluand can be found in the ToR in CS4 and CS5. In CS4 it is highlighted that no “extra-resources” are going to be made available and that recommendations requiring resources should be avoided or if suggested would need to be listed in order of priority.

Regarding the authorship of the ToR, since metaevaluation findings from CS6’s ToR indicate that these were part of a Review Notice sent by university authorities to program staff, it may be claimed that generally ToR are authored by the commissioners. However, while the lack of the actual Review Notice prevents us from knowing whether any names are identified as the authors of CS6’s ToR, the explanation of the meeting with the Dean in CS6’s Introduction announcing the closure of the program and perceived as the main trigger of the review helps to connect the ToR with this particular individual. In the verbal data (A) blames this Dean for eventually closing down the program, thus suggesting a personal factor involved:

A208. This was er the dean] who actually shut down the program so that dean’s name is (P22), he is doing the faculty so—

Nevertheless an actual quote in the report by the same Dean expressing that

“[University initials] has already reviewed other aspects of [L6] Language in previous reviews (4 in total) and so we have all the information we need regarding other aspects of viability”

clearly indicates that the dean perceives himself as a member of the top university administration with authority to also speak on behalf of the university. As a matter of fact, (A) assigns the Dean the role of deliverer of higher orders indicating, in passing,

unprofessional practices:

A206. Er this particular dean, the new dean, his role is do what vice chancellor tells him to do, regardless of data.

Similarly, the instructional character and the verbal third person used in the ToRs of CS2, CS4, and CS5 lead to claim that these may have been written by the review commissioners⁴⁶. However, only impersonal references to sponsors are found. For example, the Faculty and University’s Strategic Plans (CS2); “The School” and “the College of Arts and Social Sciences” conducting “a series of academic reviews”, and more specifically “the Dean of the College” to whom the panel “will report to” and whose name is not identified (CS5); the review being “offered by the Faculty” suggesting that it is the administrators in this Faculty who are behind the writing of the ToR (CS4). Interestingly, in CS4’s “Review Process” chapter, authorship of the ToR is attributed to a “Pro-Vice Chancellor (Quality and Learning)” who “finalised the ToR” “following consultation with the Dean of the Faculty of Humanities and Social Sciences” where no specific names are displayed.

In CS1, as mentioned earlier, the Head of School is identified as the commissioner and appears in the “Brief and Scope” section of the report. In CS3, no reference is found but as a document attached to CS2 suggests the same commission that it is implied from CS2, that is, university management. As for CS7, as mentioned, the title of the document

suggests a policy-driven review which is confirmed through interview data revealing mandated cycles of review established at this university.

3.3.5. Terms of reference: Membership of the Review Panel (RP)

A final section found in CS2, CS4 and CS5's ToR is the membership of the RP. However, the level of detail and consistency of the information presented varies. For example, in CS2, while all 4 names of Professors are provided, the membership of the two external reviewers reveals only the University but not the Faculty or School they are affiliated to. It is only through an online search that these two members can be associated with languages. Conversely, the two other reviewers' membership reveals in one case the Faculty to which the reviewer, also chair of the panel, is affiliated to ("Engineering") and in the other case the administrative role this reviewer holds in a "Centre" in connection with "Asia". However, their internal university affiliation is only discovered later in the Introduction section since no University name is added to these details. Therefore, this panel shows two (2) internal members of the university and two (2) external with a background in languages.

Further information about the panel is revealed though in the Introduction section. For example, that the composition of the panel was "constituted by the Dean of the Faculty of Arts and Social Sciences" without identifying the name of the dean which therefore reveals that commissioners not only produce the ToR but also select the reviewers. One interesting finding in this description is the characterization of one member as "an internal though independent chair". But what exactly makes this internal reviewer and chair of the panel *independent*? Does this statement imply the non-independence of the other internal reviewer? This characterization seems to be an important one to highlight in the report since a similar reference is found in the ES stating that this review was conducted "by an

independent panel”. But, why the need to emphasize the concept of ‘independence’? What exactly does “independent” mean in this context? and how can “independence” be put to the test or proved if it is so important? Have any measures been taken to ensure the panel’s independence? Perhaps, independence would be clearer with an explanation about the process for selecting reviewers, that is, about the selection criteria, the means used to contact the reviewers, the pool of candidates contacted, the number of candidates who had accepted, etc.

In CS4, interestingly, while a description about the selection process is provided as an introduction to the composition of the RP, what is learned is unclear and incomplete. To begin with, according to the writer, the process involved the development of a list of potential candidates “in consultation with” the Pro Vice-Chancellor but it does not explain who was involved in the consultation. However, this information contradicts what is found in the “Review Process” section of the report where the writer states, that “the Pro-Vice Chancellor [...] appointed a three-member Review Panel”. Furthermore, while it is known that three are the final candidates, it is not stated whether these were the only ones on the list, the only ones accepting, or whether these were selected among other ones who had also accepted. An additional relevant finding ensues from how the writer chooses to express that candidates “agreed to *serve* [emphasis mine] on the Review Panel”. Clearly the term “to serve” implies to “perform duties or services *for* (another person or organization)”⁴⁷. This seems to be a relevant point as the following quotes from three different interviewees who participate as experienced reviewers in Australian universities show how reviewers are not remunerated for participating in Review Panels. In (T)’s view, the fact that it is “unpaid” is contrasted with the “enormous amount of work”, thus signalling a perception of unfairness⁴⁸.

O26. [...] and, you know, these are senior people often, you know, they don't normally charge or get paid honoraria for doing it. You just do it as part of your work um and, in no case— [...].

R34. Well, that's where,) yeah, no, it's not remunerated.

T56. Er also, we need to remember that the job of sitting in one of these panels is very much a freelance kind of job. You don't get a remuneration for it, and I can tell you that the number of hours that are going to sit in one of these panels, and the level of, well, personal involvement that goes with it, sometimes er intense interpersonal relations, is very high, I mean, is an enormous amount of work, which is unpaid. [...]

Regarding the members of the panel themselves, the role of "convenor" in the RP is indicated which seems to be used as an equivalent of the "chair" role assigned to one of the review panel members in CS2. However, in neither CS2 nor in CS4, this role is described.

What is characteristic in CS4's section is that not only are the names and academic titles of the three (3) members of the RP provided but also a brief CV descriptive paragraph with details such as their undergraduate and graduate studies, areas of research interest, teaching, current positions, awards, etc. These descriptions reveal various findings. First, that all three hold the academic rank of professor, and second, that two are from outside the university while one is from inside. The internal member, in the role of "convenor", has no affiliation to languages, and of the two external members, one is affiliated with languages whereas the second one's background is characterized as "eclectic". However, a careful reading of this last reviewer's biographical details shows expertise in Law, in language teaching, particularly, "reading legal materials" in a specific Asian language, and in teaching and researching "insolvency"⁴⁹.

Interviewee C in confirming the composition of the Review Panel where he identified himself as one of the members

C62. [...] The review panel was constituted of 1 person from the (U3) who was not at all in the area of languages, and then, 2 outside people both from languages but 1 from an (AL) area and 1 from a (EL) area, I being the (EL) area person.

Q63. So, 3.

C64. 3 people on the panel.

Also stresses that the presence of external reviewers with language affiliation is one of the characteristics that makes CS4 qualify as a “good process”:

C75. [...] But I think that what was good about that process was that you had representation, um external representation of *experienced*— external representation of the languages areas. I think that’s, that’s very important.

In CS5, the “Membership of the Panel” following the ToR does not provide any description and consists of a list showing the first reviewer with the role of panel chair and four (4) additional members. Interestingly, an additional individual appears in this section identified as the “Secretariat” which adds to a total of six (6) members.

Some details, such as the names of the members of the panel have been covered with black ink⁵⁰, and while their university role is mostly not covered, the university role of the third internal member is either covered or it needs to be assumed that this member is an academic with no further roles. However, (M) who identifies himself as a member of this RP (“I’m not quite sure what the process for choosing *us* was”), states that “there was a higher dean like a PVC [Pro Vice Chancellor], or something”, and since none of the rest of the members reflect this role, this may very well be the role of the third internal reviewer. But, if so, why would be important not to reveal this role? Or is it simply an arbitrary decision?

Furthermore, from the uncovered information it may be learned that three (3) of these members are internal to the University, while two (2) are external. (I) verifies the variety of internal and external reviewers as an expected characteristic of the panel:

I4. [...] and as well, there was, of course, the panel composed of internal members of the university and external er reviewers while (M) justifies the selection of such high number of panel members with the size of the evaluand:

M8. [...] because it was the whole school of languages: Linguistics and the three or four five languages that were represented, they had to have a wide spectrum of evaluators, [...]

Thus, on the one hand, from the two external reviewers, one is a Professor with a background in a European language who holds a prestigious academic title, and the other holds the position of “Dean of Graduate Research School”. On the other hand, the university role of the three internal reviewers reveals a “director of a research school” who in turn is the panel “*Chair*”; a “Dean of education” and the third one a member associated with related School(s).

Therefore, the fact that the number of individuals in primarily management positions seem to exceed those with mainly academic responsibilities shows an administrative orientation rather than educational in the panel. Furthermore, the “secretariat”’s membership is also provided revealing “Executive Officer of Education Projects and Planning”, thus denoting a further member of the administrative staff of the College, though the name of the University is not included in this case. But what exactly does a “secretariat” do in a Review Panel? What does his or her task involve?

Moreover, this would increase the total of internal members in the Review Panel to (4) in comparison to two (2) external. However, the real imbalance comes more from the unevenness in number of members with languages affiliation. Hence, among the internal ones, the one with a covered up role is affiliated to a School where languages are taught situated in the “College of Asia and the Pacific”, and the external member with a prestigious title is also affiliated to a School of languages (which through verbal data can be identified as (M)⁵¹). The other three members have no background in languages which explains why (M) finds the representation of languages expertise in the panel insufficient in comparison to the size of the panel.

M8. [...] er and that one spectrum was not a concentration of er language specialists. It was, you know, somebody from the romance languages, and somebody from linguistics, and] do you see?

In fact, while (M) considers one of these members as a linguist, the linguistic background seems to be overridden by the fact that this individual is perceived to have come to the review as an administrator. Therefore, (M) is inclined to conclude that the professional expertise most closely connected to languages is (M)'s own expertise in "curriculum" and therefore the member of the panel with the most legitimacy to contribute an academic perspective to the review:

M10. Um there was somebody, I've forgotten his] I think he's a linguist as well. So there might have been one, one linguist and one er applied linguist but he was also a dean in another university, and I was the person who knew most about curriculum in language departments, [...]

Interestingly, while recognizing a lack of awareness about the selection process for reviewers as mentioned, (M) thinks that the primary reason for having been selected was not necessarily expertise but something else (M) decides not to mention:

M10. [...] um but I'd probably been chosen because I was in fact-

One possibility may be the prestigious award reflected in the report (and not covered up) which may contribute to reinforce (M)'s professional entitlement to judging the program.

According to (I) there seemed to have been a consultation process for the external members who are the ones "selected to be in the panel" in contrast with the "de facto" reviewers who "were the members of the, of our research school and, and so on. This was the people, the internal people". However, the consultation seems to have been badly managed since (I) concludes that "it was a very confuse [confusing] affair". First, (I) complains that it "was not wide enough", second that "some of them were appointed

without consultation”, and third that the appointment of a “language teacher” was done in a rush “by the Head of school”, and only when

192. [...] it was very clear that there was no language teacher in the panel, then there was a reaction and they had to include someone but they include someone who was not adequate.

But what does (I) mean by ‘inadequate’? The following quotes show the

characteristics that in (I)’s view disqualifies (M), in particular, as an adequate reviewer⁵².

Hence, (M) is described by (I) as a non-vocational “language teacher”, research-oriented and with limited practice and knowledge in application within the discipline:

17. [...] the main difficulty with the panel was that there was only one person that could be considered a language teacher but I know this teacher— this person well, and it was one of the typical intellectuals who didn’t find a job in an English department or a philosophy department or something like that, and ended up in a language department but mainly to do research in an area. Er this person was characteristic of this kind of people who essentially can speak the language but it’s— they don’t have as a— in their heart to be language teacher.

1110. [...] the only, the only language teacher was someone that I know and who, as I said before, was someone who never taught elementary courses because it was, it was— this person was incapable of doing it, only taught at the advance level courses when the students have already the knowledge of the language. [...] All this is a view that it’s completely different from someone who only up, and only deals with the intellectual issues because there is no language barrier.

As for the other members of the panel, or better said the panel as a whole, (I) explains how the lack of knowledge about language teaching raised tensions in the discussions and even contributed to ignoring claims by the appointed staff representatives:

17. [...] So we have a whole panel of people that we have to make, I mean, to try to make them understand some of the limitations of the, of the discipline and it was a very dramatic event, I mean, in the sense of er a lot of misunderstandings.

174. Well, we were, we were all the time involved because, I mean, we had these consultations, we had these meetings, we had, then, the forums, we had— so the, the, the difficulty was not if you were participating or not, the problem was to be heard.

176. Well, no, they engage with the program staff. The problem was that um I mean er they would not listen

Moreover, the difficulties mentioned by (I) above about the lack of knowledge combined with the other university “top down pressures” associated with reducing resources are perceived as “a constant battle” and which reviewers fail to understand:

I116. [...] And so it's, this is, this is the core of the, of the, of the (xx) of reviews, is that people that review are ignorant, in the well sense of the term, in the sense that they don't understand the issues, and—[...]

Interestingly, (M) *does* seem to be aware about contextual disadvantages such as the push to make language teaching programs more cost-effective, which is also perceived as a general trend across the Australian university system:

M8. [...] Er this has to be set against the general fact in Australian universities that universities do not understand that teaching and learning a language isn't just a skill that can be sort of siphoned off and er at sort of basic competence, okay?, and they don't, they also don't understand why it can't be done more cheaply.

And, in particular, the program's unfavorable position in the university, and the pressure and lack of institutional support:

M8. [...] So, they were not part of the university in er-part of the research, part of the university and therefore any research that they did was kind of a bonus, but-all right? So that-but that's another piece of background.

M50. [...] in a department like the (U14) they're under terrible pressure, they are not cohesive. These things are blindingly obvious.

M54. They never felt supported, they never felt supported, nor in (S/T2). They had never felt supported. That's, that's the real issue. [...] and that was a mistake at the (U14), that they didn't feel supported institutionally

(M)'s awareness of the lack of language teaching expertise in the panel except his own, and the program's lack of support by the University lead (M) to make the conscious decision to adopt, on the one hand, the role of a language expert to clarify doubts raised by other members

M10. [...] I made it my business to try and make positive suggestions and so, well, this is, you know, [...] and also That was my job: I took that as my job. So they would say

‘what do they think they were teaching, culture?’ and I would then make it my business to explain that.

and on the other, the role of an advocator for the program (“to get committee on side with the languages.”). Interestingly, (M) takes pride in the success not only of making the internal members of the panel “to come along with me”, but particularly of helping a perceived University high authority whom (M) asserts with conviction was the person behind the decision for choosing (M) as a reviewer. The gratification of such an experience makes (M) assess the work done by the Review Panel a success:

M10. [...] And I And I must say there were some members of the committee who knew, but there were other commi-particularly the dean who really didn’t understand, but she came along with me and then there was a higher dean like a PVC or something and he was terrific er in his] he was the one who’d, I think, chosen me by all accounts and he was-he just did not understand what they were doing but he came. So the committee, I thought, worked tremendously well, felt the committee had worked extremely well and we did a repor-

Q11. -Cooperativ[ely.

M12. Very] cooperatively.

As for information about the reviewers/Member of a RP in the rest of the reports (i.e. CS1, CS3, and CS6), in CS6 apart from information on the TP where the two (2) names and academic titles (“Dr.”) of the submitters of the report, and therefore of the conductors of the review, reveal two (2) program members, in the Introduction section, one of them is identified as the Head of the language area under review⁵³. Furthermore, as mentioned, the composition of the RP to which CS6 is to be submitted is also provided which reveals the names of the three (3) members with their university roles and/or academic titles. However, no association with any particular university is made explicit but the indication of “(External chair)” next to one of the members implies that the other two (2) members are internal. It is through an on-line search that it is learned that the external reviewer is an Emeritus professor with no affiliation to languages, and that one of the internal reviewers is the Dean

in the Faculty of Business, and the other the Associate Dean of Education, thus, with no affiliation to languages.

As a matter of fact, interviewee (C) and (A) provide some evidence denoting a negative perception about the composition of this RP (CS6). Besides (C)'s criticism of the external membership of this panel but most of all of the lack of language representation, in contrast with CS4's RP, emphatically stressed as an important pitfall of the review:

C75. [...] For instance, at the (U4) where, I mean, in this review that they did of the (L6) program, there was no representation of language at all, and I don't think anyone external or there might've been someone external but it was not somebody to do with language [...].

(A) complains about the commissioners' self-interested selection to implement the predetermined decision to close down the program:

A12. Er it's] someone from faculty. It's, it's supposed to be from er one internal faculty back then dean— associate dean of education and er in different de- dean of different faculty and another person is er bringing from the external. She is from— someone from (C3) actually um was a mates of dean. So it was kind of— it's er committee made to make his decision happen.

In CS1, similarly to CS6, the names and membership presented on the TP as the authors of the report reveal two (2) members of the specific Linguistics Department, one a "Dr." and the second one with an MA degree who in other sections of the report can also be identified as the conductors of the review ("We were tasked to [...]" "We focused on [...]", "we included students [...]", etc.). Therefore, they are internal to the university, and to a certain degree external to the program since they both live under the same School of Languages. It is interesting to note how these reviewers identify themselves in the Methodology chapter as "Investigators" instead of 'reviewers/evaluators'.

Regarding CS3, though embedded in CS2 as a "Self-review", this report provides details neither of the self-reviewers nor of the RP to whom this report is to be submitted in

contrast with CS6⁵⁴. And while the use of the first person subject pronoun “we” and the possessive adjective “our” seems to point out that internal members of the School may be the conductors of this review and the authors of the report, the overall impersonal mode in describing findings and the lack of references to methodological techniques used seems to indicate that it was a single internal individual doing the review and writing the report on behalf of the School. These conclusions are evidenced in expressions such as “The School’s intention is to [...]”, “The School is committed to [...]”, or others describing what the School ‘has’ or ‘offers’ (“We have partnerships with”, “we have a relatively low turnaround”, “we offer a larger number of majors”); or demonstrating activities that have been carried out (“we have as a group also worked hard”, “We had a Skype presentation”, “we have actively participated in”); or when referring to the School as a whole (“our School”) as well as to the different constituencies and layers of the School (“our disciplines”, “our language programs”, “our external funding sources”).

In CS7, while no names of reviewers are found in the report, as in CS3, however, in contrast, many of the grammatical subject references expressed with the personal pronoun “we” are in connection to data analysis, thus suggesting a collaborative gathering and studying of these data. Examples of these are “we do not have data that would suggest an explanation for this” or “we currently do not have a sufficiently strong cohort [...]” but the clearest one is “As a result of this review we have undertaken an analysis of [...]”. As a matter of fact, oral data by (R) confirms that “The teaching team, all of those involved, are part of the evaluation” and explains how different individuals, (R) included, were in charge of different tasks which, involved the writing of the report too:

R10. [...] What I did do, my part, the part that I gave to me was to review, to conduct focus groups: focus groups with students, focus group with the staff and focus group with external, [people external]

Q11. Stakeholders.]

R12. [...] Now, others did comparisons with other universities, now, whether we like it or not, we are required to benchmark, it's called, with other universities, [...]. Er others er looked at what is the situation with language teaching and learning more globally, a lot of us are involved in that, but there was a kind of literature review, a context statement, so, one person was responsible for that, someone did the benchmarking, I did the focus groups, someone else looked at just the data as it stood um and the student feedback and so on, and so different people looked at different um aspects.

In conclusion, the existence of ToR in CS2, CS4 and CS5 and CS6 is clear evidence that the inclusion of ToR in the reports ensuing from evaluations of language programs in the Australian Higher Education system is common practice. As for the case studies where no ToR are found, in CS3, the similarity evidenced in the areas to be reviewed in CS3's ToC and Results chapter with the ones in CS2's ToR leads to the assumption that the same ToR apply to both reviews. In CS1, while lacking a ToR, information in the "Brief and Scope" section describing details about the commissioning of the review may lead to claim the likely existence of such a document. As for CS7, Terms Reference are lacking, possibly due to the fact that the internal policy that triggers the review provides "very precise guidelines and procedures" as interviewee (R) states though it may exist and not be made available for the purpose of this metaevaluation⁵⁵.

3.4. Evaluand and evaluation context

3.4.1. Description

According to the ERC, an introductory chapter should provide information to contextualize the evaluand and the evaluation itself. Therefore, the checkpoints to be expected in this chapter are a description of the program or phenomenon being evaluated with goals and historical context, if appropriate; the target population for the program and the relevant audiences and stakeholders for the evaluation; a review of related research,

and an overview and description of the report structure. Moreover, as mentioned, the evaluation purpose and the evaluation questions may be included here unless explained in the Methodology chapter that follows.

3.4.2. Evaluand, Program Goals and Objectives

The reports with an explicit Introduction section are CS2, CS3, CS4, and CS6. However, while CS6's Introduction is independent, CS2, CS3, CS4 have it integrated within the Results chapter, and similarly, CS7 also offers introductory information within the Results chapter even though there is no section identified as such. Moreover, CS1's "Brief and Scope" presents background information which together with the Program description that follows qualify as introductory material. Lastly, CS5's only introductory details consist of an independent section with a description of the program.

While CS1 and CS5 have independent sections for this checkpoint, in CS2, CS3, CS4, and CS7 it is intertwined with introductory material introducing the Results chapter. As for CS6, as a detailed program description is included in CS6' ES, there is not one in this section.

In these case studies, program descriptions where goals are explained are found in CS1, CS3, CS6, and CS7. However, a closer look at all the descriptions, with or without goals, reveals similar emphases. For example, in CS7, besides the tabled "Program Summary" showing the name of the evaluand, and other technical and administrative details ("Code", "Division", the "School" it is linked to, "Year first offered", "Duration of the program"; and "Total unit value"), which denote a policy format, a description with program goals is presented in the Results chapter. Interestingly rather than describing the objectives of the program, a table is presented showing how the program's objectives have been shaped according to 7 "Graduate Qualities" that students should leave the program with, thus,

demonstrating efforts to align program's objectives with specific predetermined strategic goals.

In CS6, similarly, the program description reveals an emphasis on the program's role in a) "attracting" students both local and international; b) its alignment with both University and International policies; c) its responsiveness and adaptability to changes "over a short time"; d) its active outreach engagement both nationally and internationally; e) its proactive engagement in marketing; and f) its importance as the "main driver" for exchange in the geographical region of the second language the program teaches. Additionally, however, emphasis is also on the program's contribution beyond "student numbers and status as a service area", thus expressing an underlying defensiveness from a critical view towards focusing the review on "finances"⁵⁶, and at the same time expressing willingness to further contribute to increasing student numbers ("significant" opportunities "for growth in student numbers and enhancement"). In sum, it could be claimed that CS6's program description reveals the program's compliance with a set of priorities, particularly with increasing enrolments. Moreover, the fact that an exact copy of this description is used in the Conclusion section of the report reveals an interest in highlighting further these same emphases, thus, implying in turn the high stakes of this review.

In CS1, the program description after the "Brief and Scope" section contains general descriptive details in the first paragraph about the program such as number of points the students may attain, length of the program, range of languages offered, and the programs' general aims. However, the two remaining paragraphs focus mainly on the topic of enrolments, specifically on the number of enrolled "fee-paying students" versus "sponsored places" both overseas and Australian, and also on the degree of popularity of the languages in terms of their "enrolments" during a specific four-year period. Therefore, the focus of this

description falls mainly on student numbers in line with the stated main concern of the review (discontinuations), and particularly with the reviewers' stated additional purpose of the review to also "increase enrollments".

In CS3's program description⁵⁷, besides the description of the School's strategic objectives mentioned earlier to be aligned with "the University and Faculty strategic plans", structural changes, both at macro and micro levels, and staffing and student/staff ratios are described. For example, the writer highlights first of all the reconfiguration "from a departmental to a School model in order to secure administrative efficiencies at scale" which he/she finds a justification through alluding to a broader existing trend in "many other faculties in Australasia"⁵⁸. Moreover, at a micro level, the development of programs and departments is described in terms of additions, eliminations, splits, renaming, moving, becoming, joining, developing into, and so on. The section on development of staffing and student/staff ratios data describes percentages resulting from benchmarking staff/student ratios between the School and the Faculty, across departments showing the School's higher ratios as a result of a reduction in staffing except for those programs which "contribute units into Programs". Therefore, compliance with a range of University priorities, stressing a general efficiency trend and its application in this particular School showing some deteriorating consequences are the main themes identified in this description.

In CS4's description, while the purpose of the program is tackled together with other historical details such as the year it was introduced, the majority of the information is of an administrative nature such as the number of units needed to complete the program, the student population the program targets, the two different language programs offered, the "adjunct" nature of the program, and the range of languages taught both at the university and cross-institutionally.

CS2, and CS5's descriptions share with CS3's a focus on structural changes though, unlike CS3, these lack an explanation. For example, CS2's description consists fundamentally of a listing of the structural changes that it has undergone since "the Faculty of Arts moved from a departmental to a school model". In CS5, while the name of the section suggests an aim to provide a historic account of the School ("The School of Language Studies – A Brief History"), no information is contributed about the educational dimension/mission of the School, and the primary focus is on additions, creations, establishments, restructuring, appointments, abolitions, hiring freeze, and reductions. But, why focus the history of the School on the theme of restructuring? Why is it so important? Are these changes perceived as part of strategic priorities and therefore used to demonstrate compliance or may the review policy require the inclusion of this information?

The section ends on a positive note highlighting first the "less material impact" on the School by the most recent change (from a "Faculty" to a "College"), and second, that the "only change of note" has been the creation of a "centre" within the School. May this respond to a need to evidence the School's willingness to supporting efficiency-related University goals, particularly, restructuring?

3.4.3. Evaluation Audiences and Target Population

In these reports, this information can be found scattered across sections. In CS1's "Brief and Scope", 2 Heads of School are the first audience identified as the commissioners of the review, one in continuation of the other, as well as a "University Academics Programs Office" and the name of the individual responsible for the previous study from which the current one follows up. Therefore, it could be claimed that the main audience for this review may be University administrators. Moreover, while "students" are central in

resourcing discussions, program staff are also identified as a target group in the ES, Methods chapter and Results chapter. Interestingly, the fact that the recommendations are divided into both academic and administrative categories implies that administrators may be considered to some extent as target population in this review, too.

In CS2, CS3, CS4, and CS5, a similar pattern may be identified in terms of review audiences. On the one hand, the administration and therefore the commissioners such as the University and the Faculty (CS2), the University, the Faculty and the School (CS3), the Faculty of Humanities and Social Sciences and the University (CS4) and the College of Arts and Social Sciences (CS5) are presented as the first stakeholder group, and therefore these may represent the main audience for the review. The clearest illustration of main audience is for example the instruction in CS5's ToR asking the reviewers "to report to" a specific "Dean".

Regarding target population, in all four reports, the same variety of groups which are consulted can be identified (e.g. students, staff, other university members and external stakeholders). However, in the Results chapter, the most cited group is "The School" with the exception of CS4 where the University and the Faculty are often addressed as stakeholder groups accountable for languages. In CS2, for example, the School is referred either as an entity ("[...] its strategic directions [...]") or personalized ("The School [...] generally understands the need to [...]"), or sometimes also referred to as "members of SXX", "SXX staff" or any other unit within the School such as Departments, Studies or Programs. In fact, the interview schedule for CS2 in Appendix 1 reveals that thirty-three (33) members of the university from the School and from outside the School participated in meetings in contrast with eight (8) students who were interviewed. Thus, this rate shows

that the target population and therefore the priority in this review is to learn about perspectives from staff. This suggests a degree of educative focus in this CS.

In CS3, included in CS2's appendices as documentation to be reviewed in CS2, the review audience in CS2 applies to CS3 (i.e. upper University authorities), and as far as target audiences, findings have to be inferred from evidence in the Results chapter. Hence, various groups and individuals are mentioned such as "Pro-dean Teaching and Learning", College Board, "The Dean", and "External stakeholders", "exchange-study partners", "language specialists", "academics", "honorarys", "students", "casuals", "administrative staff".

However, while these groups reveal a mix of potential target audiences, "The School" is the most cited one in this self-review report too. Moreover, data in relation to these groups is used to demonstrate compliance or show managers' satisfaction with the School's fulfillment of priorities (e.g. anecdotal comments by a manager commending the School; supplied data on students in relation with high response rates and student satisfaction and on staff with the conduct of Departmental reviews). Therefore, it is possible to claim that it is the main administrator leading the School, the individual accountable for how the School is meeting broader goals, thus, who is the main target population for CS2.

In CS5, similarly to CS2 and CS4, the target groups mentioned are "the School, academic staff, students and external stakeholders" and within these groups, the Head of School, whose name is covered in black ink, and "many of the staff of the School". Among students "undergraduate coursework and higher degree by research students" are specified as participants, and the particular external stakeholders too (e.g. embassies and a range of cultural organizations as well as the "state and federal governments"). Additionally, in the Results chapter, areas from within the University are also mentioned such as "College of Law", "College of Asia and the Pacific", "School of Music", etc. organizations such as "The

Australian Academy of the Humanities”, as well as other universities (“Charles Darwin University” or “Group of Eight Universities”). Nevertheless, “The School” is the most cited group of all which is further supported by evidence in the recommendations revealing that all the work suggested is to be undertaken by “the School”.

In CS4, while the anonymous writer expresses in the ES that the RP examined the program’s “target audience”, this is not identified. However, the ToR reveal a variety of populations the reviewers are to collect views from (“current staff, students and external stakeholders, including employers, alumni and representatives of relevant professions and other relevant academic areas of the University”) which is also reflected in the “Review Process” section and in the Results chapter. Similarly to CS2, this diversity is evidenced in Appendices A, B and D, showing a variety of staff (e.g. casual staff, teaching staff, the Asian language staff, staff from the Faculty of Humanities and Social Sciences), a variety of students (e.g. international students, all students from across the University, students in languages, students from the two other [regional] universities), and external stakeholders (e.g. high school career advisers, other Australian universities, and a state educational agency). However, the fact that in CS4’s recommendations actions are suggested to the two Schools, and the language courses⁵⁹ but also to the University, and the Faculty lead one to claim that these also as target population for this review, as in CS1.

In CS6, on the one hand, those involved in the commissioning and management of the primary review that is, the University, the Faculty of Arts and Design, specifically the Dean, and the Review Panel to whom CS6 is to be submitted, may be considered the main audience for the review. On the other hand, CS6’s program staff managing the program, in this case, the reviewers who include the Head of the program and an additional staff member, can be considered the primary review’s target population since they are the staff

accountable for the program. However, CS6's Results chapter and Appendices data collected in relation to a varied range of students are presented (e.g. higher unit students, inbound, outbound students, internship students, exchange students, UG and PG students, students from across campus, Chinese international students, alumni, mentors, student club), as well as staff (e.g. continuing staff, casual staff, full time staff, languages staff, "the staff of the [program under review] language area"), and external stakeholders (e.g. Government from program's taught second language native country, newspapers, local Universities, exchange partner universities, language learners from Schools, a High School), and therefore can be considered target audiences in this self-review.

Likewise, in CS7, the University, the Division and the School are the first groups of stakeholders mentioned in the introductory information within the Results chapter, thus indicating the main audience for the review. Moreover, as stated in the report, the "focus" of the review, and therefore the target population are the students "enrolled in the program" which does not include those who come from "a broad range of programs". Among these students are international students, ATSI⁶⁰ students, NESB⁶¹ students, students with CALD⁶² backgrounds, rural students, students in the Disability category, and Graduates (employed and HDR⁶³). As for staff, the teaching team is mentioned often as the group leading the program as well as the group responsible for the conduct of the review. Therefore, while the target population in the analysis is clearly students, "the program team" may be considered the object of the review, which is clearly supported by the recommendations since all of these make reference to actions to be undertaken by program staff. However, in line with findings from the previous reports, apart from students, other stakeholder groups taken into consideration are other areas of the University (e.g. Division Marketing, `Career Counsellors, and Career Services, the Library, etc.) as well as several

internal and external entities and organizations (e.g. the School sector, the Chamber of Commerce, the Australian Public Service Commission, the School of Languages, the Modern Language Teaching Association, four partner universities local and from overseas, among others).

3.4.4. Review of Related Research

Some reports such as CS1, CS6 and CS7 have some references to related research while others such as CS2, CS3, and CS4 don't, and in CS5, possible references are identified but they are only partially addressed. Moreover, even when references are clearly brought up, their sources are not always made explicit and they all appear intertwined with review findings.

For example, in CS1, while discussing the external factors that may affect discontinuation of students, the concept of "good discontinuation" is introduced as "another point sometimes raised with discontinuation" but without quoting sources. Similarly, no sources are provided when related research is used to justify student responses, for example, on student perceptions about the program ("in line with previous research, learning languages also brought vocational and personal development and increased students' cultural and linguistic sensitivity"). More importantly, although the existence of a previous quantitative study is made explicit, this study is not described nor are its conclusions explained apart from illustrating the study's percentages on *when* students discontinue.

In CS6, when addressing attrition for example, the authors of the report provide explanations about different kinds of attrition, and past results from their own research on

attrition in relation to their program. Moreover, outcomes from a study on attrition in Australia's higher education system are presented, as the following quote reveals:

A relatively high attrition rate is also commonly observed in many language programs across the nation, and particularly in [L6]: this is well documented in the research literature (see Nettelbeck et al 2007, 2009, Martin and Jansen 2011).⁶⁴

However, it must be noted that while in CS1 and CS6 actual research related to what is being evaluated is addressed, in CS7, the reporter uses research references to support, for example, the role of languages in education:

In particular, languages are increasingly regarded as assets in terms of access to: knowledge, expertise and leadership (intellectual and other capital). They are regarded as investments in engagement with global markets and economies for individuals and nations (e.g. Grin 2001, 2003, 2005, 2008, 2009; Extra and Gorter 2008; Eaton 2010).

And while sources are provided in this instance in CS7, however, sources are not always indicated:

There is much international research which suggests that monolingual speakers of English are disadvantaged in the globalised workplace compared with those who speak English and another of the global languages of wider communication.

In CS5, a theoretical reference is vaguely suggested when commenting on the existence of "contemporary international debates about the nexus between language and culture" pointing out the "language-making capacity of culture and the culture-making capacities of language" without citing any particular source. Interestingly, the one research reference found and clearly identified in the report as one "commissioned by" a particular Australian organization regarding retention, is not commented on at all ("The Panel refers the School to the recent report on [...]") and its source is partially cited.

An interesting additional finding that emerges from metaevaluating related research is that while evidence is found suggesting a quality assurance review cycle triggering CS4, CS5 and CS7⁶⁵, the only references to previous reviews appear in CS1 and CS6. Therefore, this

evidence suggests that the cyclical character in reviews in the Australian Higher Education system does not necessarily imply ongoing development.

3.4.5. The Structure of the Report

Of the reports with a designated Introduction section/chapter (CS2, CS3, CS4 and CS6) only CS2 and CS6 dedicate a few sentences describing the structure of the report, that is, of the Results chapter that follows. However, CS1's "Findings" chapter offers a brief introductory paragraph which provides details about the organization of findings which entails a two-part division into findings based on the quantitative data and findings from qualitative data. Furthermore, inside the Results chapter of CS1 as well as CS6, transitions within sections are also explained so that the logic of the discussion is followed with ease. For example, in CS1, the discussion on the limitations of quantitative analysis helps to transition smoothly to the next part which asks further questions about discontinuation to gain a deeper insight into what can cause the "around 60%" discontinuation rate. The quote below shows a similar example from CS6 Results chapter:

Having discussed the strategies we have implemented in response to the past reviews and our own analysis, now we turn to our actual student numbers.

In CS2, however, an issue emerges when the structure of the report is addressed.

While references in the ToR and the ES in the third person such as "the review panel should:"; the panel identified [...]" ; "the panel agreed [...]" , etc. leads one to think that the reviewers are not the authors of the report, in the Introduction there is an attempt by the writer to establish a close connection of the Results chapter with the Review Panel stating that "The panel *has chosen* to structure *this report* by [emphasis mine] [...]" . However, a close analysis of its contents indicates that "the panel" becomes once more a third person

thus revealing that the reviewers may not be the only creators of the discussion. Therefore, the results presented constitute a summary of the reviewers' findings and recommendations as illustrated in the very first sentence of the section stating that "The panel spent considerable time considering [...]". At the same time, while this obvious evidence is not acknowledged and the identity of the real author of this report is maintained in anonymity, this illustration also implies that the writer of this summary may have actually witnessed the reviewers on-hand tasks⁶⁶.

3.5. The Methodology Chapter

3.5.1. Description

According to ERC guidelines, an evaluation report should include a methodology chapter, whose purpose is to inform the reader of the formal theoretical dimension of the evaluation as well as its practical side. This is necessary in order to ensure a rigorous and theoretically informed methodology which can produce findings that are valid and credible. Therefore, the issues in this section of the report may be expected to deal with the evaluation approach or model being used as well as the rationale for choosing it; the design of the evaluation including sample sizes and timing of data collection; the methods of data collection, including description of data collection instruments; sources of information and data; and limitations of the evaluation as related to methods, data sources, potential sources of bias, etc. Finally, as mentioned, the purpose and evaluation questions would be placed within this chapter unless already provided earlier in the Introduction chapter.

3.5.2. Findings

In this metaevaluation study, the only report that provides a differentiated Methodology chapter is CS1 which focuses on the design of the study, within which two parts may be differentiated.

In the first one entitled “Investigators”⁶⁷, besides the identities and membership of the researchers, the specific data collection tasks that were taken up by each of the two reviewers are reflected (staff interviews and compilation of quantitative data taken up by one and student interviews by the other). The second part of CS1’s Methodology deals with data collection methods but it also includes a description of instruments, samples, and timing. Moreover, some limitations of the study are also discussed.

The description first explains the L2 languages selected from the program for the purpose of the study and the criterion for choosing them; the “Project Phases” of data collection including number of interviews conducted in each one and their nature (the first one “exploratory” thus allowing the identification of “major issues” leading to a “preliminary report”, and the second one aggregation of data and writing of final report). The timing of these phases is not included here but appears when discussing the procedure for collecting student interview data later on. Moreover, the description also includes two sources of quantitative data used, (the report from the previous quantitative study and data from a “student administration database”) and the sources and instruments for collecting qualitative data, that is, “Interviews” which includes a detailed description of “Participants” (Students and Staff) and “Procedure”. In the students interviews section, the reviewers explain the “selection criterion” and the rationale behind it (e.g. “ensure a wide variety of experiences and opinions” and variables chosen on the basis of those that characterize the “greater population” of the program); the distribution and level of representativeness

(achieved by balancing participant numbers across a number of variables such as “degree”, “entry level”, “language”, [program name] status, “age” and “sex”); and the total number of student participants. Two tables provided in this section clarify the distribution of numbers of students selected according to the variables. Furthermore, the Procedure section describes the process of interviewing and includes information such as the interview style chosen (individual or group) and its rationale; what guided the interview; the focus of interview questions; the systematic steps taken before the interviews to ensure an ethical process; the interview location; the interview length; instrument for oral data collection and the variety of interview modes (face-to-face, etc.). The main purpose for all of these steps being, the report states, the seeking of “detailed” information.

The section on Staff is considerably shorter though justified by explaining that the interviews with staff were conducted “very similarly” to the ones with students, and covers fundamentally the same areas as the ones mentioned above. The total number of interviewed staff is shown which, relative to the number of students, shows, as mentioned, a priority to focus on students’ perspectives.

The limitations of methods are also expressed through comments such as “For reasons of efficiency and practicality, only current university students were approached” or “participant numbers for each language *roughly* [emphasis mine] reflect the size of each language cohort in the program”. Moreover, in terms of potential sources of bias, a comment is made about how “Participants self-selected to some extent by making themselves available, possibly because they held strong views about their language studies” which reviewers tackle by reaching students with “a wide variety of experiences”.

Regarding limitations of the review, the reviewers include a relevant comment in the Recommendations chapter recognizing the difficulty to “project what level of

discontinuation would be reasonable [...]” due to the fact that “no data is available about discontinuation rates from other degrees and diplomas”.

As for the evaluation model, while missing from the Methodology section, the “Brief and Scope” reveals that this study is conceived as “a qualitatively oriented in-depth follow-up study” and the rationale is explained in the Results chapter in terms of the need to consider the limitations of “quantitative analysis” to find out the variety of reasons why discontinuations occur. Interestingly, while mostly qualitative, quantitative data are also used, thus revealing a multi-method approach.

Out of the remaining six (6) case studies, the only other example of what could be perceived as a Methodology chapter is a section in CS4 named “Review Process”. However, what is found mainly is cursory information about data collection and data sources, suggesting compliance partly with the ToR and perhaps also with a review policy where transparency about how the process has adhered to procedural review steps is emphasized. For example, the first description reveals the involvement of management in the process, specifically, “consultation” with the Dean of the Faculty before finalisation of the ToR and appointment of reviewers by a PVC (Pro-Vice Chancellor) as discussed earlier. Similarly, it also emphasizes that the perspectives of diverse groups were taken into account, as the ToR requested, the exact dates when the review took place, and a brief description of data sources and collection methods. For example, how submissions were collected via “individual invitations” to a number of university members, the involvement of a University “Quality and Reviews” office for the assembling of documentation “to assist the Review Panel”, and the list of stakeholder groups participating in the interviews with the Panel for which evidence is provided in the Appendices. However, the contents in these Appendices are enumerated rather than described. For instance, the writer mentions the “list of

invitations sent” for the submissions shown in “Appendix A”; the list of documents assembled shown in “Appendix C”, and “a copy of the schedule of the review panel visit” which appears in “Appendix D”. However, while the contents of Appendix B are also stated (“nine submissions were received”), the writer does not specify that this information can be found in Appendix B. Finally, to further emphasize transparency, the public exposure of review information such as “The terms of reference and other details about the review” in “the University website” is also stated. Lastly, this section ends with a thank you note, as noted above (“the panel wishes to record its thanks to [...]”).

Therefore, no description about how interviews were conducted, how submissions were called for, how the material collected was analysed, etc. is included. Thus, rather than a Methodology chapter, these details qualify as a synthesis of methodological information which according to the ERC is typically recorded in the ES.

In CS2 and CS5, methodological evidence consists of, similarly to CS4, cursory details on data sources and data collection methods which at times reveal contradictions, and lack of clarity, too. Moreover, since these cases lack a designated Methods chapter, findings are scattered across several sections of the report.

For example, in CS2, the methodological information that is found consists mainly of the summary provided in the ES and sheds light on data sources and collection methods. However, this information is cursory and to some extent contradictory. For example, as mentioned, while in the ToR the “School self-review documentation” is presented as the only data source to be considered by the RP, in the ES and the Introduction the writer states that the RP took into account data from “meetings” during a three day “campus visit”, and from “submissions”. Also, the fact that the constitution of the panel was decided by a Dean, as stated in the Introduction, suggests a top-down approach to the review.

Regarding submissions, unlike CS4, there is no specification as to what stakeholder groups comprised “the interested parties” from whom submissions were called, nor is there any information about how the submissions were called, or whether the three submissions received⁶⁸ corresponded to the groups of stakeholders who may have been contacted, or ‘invited’, whether they were selected among others or these were the only ones received. Of particular interest is how the submitters’ anonymity was secured, particularly, given such a low response rate. Moreover, while the report states that out of the three submissions one was “open” and two “in confidence” which may lead one to think that some kind of process was put into place, confidentiality in this case seems to be associated with the act of making its contents public, as the “open” submission in Appendix 3 entitled “non-confidential submissions” proves, and not to do with the process of acquiring/collecting submissions as also happens in CS4.

In terms of meetings, the variety of perspectives and of topics covered (“wide-ranging discussions”) on the one hand, and the quantity of stakeholders involved (“numerous stakeholders”) seem to be the most emphasized traits. The diversity of stakeholders, in fact, is mentioned twice and described, too, in the Introduction which identifies several groups “from current students to academic and professional staff, and from new and senior academics to the current and incoming Heads of School.” Furthermore, as in CS4, an interview schedule in CS2’s Appendix 1 is used as evidence supporting this description, which helps us to learn, as the table below shows, that forty-two members of the university from the School and from outside the School participated in meetings which included a variety of individuals, except for external stakeholders, particularly, twenty-four in academic/academic related roles (12 chairs; 4 directors of “Studies” areas; 4 “level E academics”; 3 “recently appointed academic staff”; and 1 School Coordinator for Teaching

and Learning), and nine (9) in managerial roles (1 the Head of School; 1 School Administration Manager; 1 Director of Strategic Planning and Operations; 2 Associate Director of Finance and a School Finance Manager; 1 Pro Dean Research and 1 School member of the Research Advisor Committee; 1 Associate Dean UG; and 1 Faculty of Arts and Social Sciences Dean). It is interesting to note an additional member of the School difficult to categorize since this person is identified as a “Diploma of Languages project owner”, and therefore does not fit either academic nor managerial⁶⁹.

Additionally, a total of eight students were also interviewed in two different sessions. What is known about these participants is that the first four are a “sample of postgraduate students” and the other four “a sample of undergraduate students”. Thus, it may be claimed that the rate of 34/8 shows that the priority in this review is to learn about the perspectives from staff (excluding casual teachers).

Participants		Languages
Managerial roles	9	
Academic/academic-related roles	24	Chairs (12 languages) Directors of “Studies” areas (4- 1 Asian) Level E academics (4-?*) Recently appointed academics (3-?) School coordinator Teaching and Learning (1-?)
Students	8	Postgraduate (4-?) Undergraduate (4-?)
Unidentifiable	1	Diploma of Languages project owner (?)
External Stakeholders	0	
	42	

Table 5. The range of interview participants and languages represented in CS2.

*indicates unknown language

Moreover, apart from dates, times, meeting places and breaks, other evaluation activities can be learned from the two-page well-presented interview schedule such as four (4) scheduled time slots allocated for the panel to discuss and/or draft the report; a walking tour through facilities with the administration manager; and one opening for meeting requests. However, the lack of an explanation in this Appendix prevents us from learning whether requests were in fact made or whether there were additional meetings with reviewers.

In terms of the limitations of the study, in the Results chapter a statement is made regarding the small number of student informants (“Whilst it is recognized that the sample was small [...]”) though no explanation is provided as to why this happened.

In sum, in terms of methodology, the report of CS2 reveals a lack of essential evaluation methodological issues suggested in the ERC and the details on the design based on data sources and collection methods are superficial. Furthermore, the fact that the incomplete methodological information in the ToR is complemented in other sections suggests the conception of this report as an internal document shared within the University administration stakeholders. Thus, it could be claimed that an administrative member of the University is behind the writing of the report, whose job it is to reflect regulatory review procedures have been followed in the review which, as seen, have little to do with methodological rigor and more with compliance. On the other hand, a wide range of stakeholders was consulted, suggesting that there was some concern for educational questions and the interests of those other than in administration.

In CS5, similarly, the summary of methodologies in the ES constitutes the main content found as far as methodology is concerned, and as in CS2 and CS4, it focuses on data

sources and data collection methods with no discussion on any other methodological issues. As findings from the ES explained above show, an emphasis is placed, similarly to CS2 and CS4, on range of stakeholders in line with the ToR's emphasis to "widely" consult stakeholders from inside and outside the university "as appropriate".

In this case study, the data sources and collection methods used are on the one hand, submissions provided to the RP (the panel "considered the submissions by [...]") and meetings with the RP, as in CS2 and CS4, and on the other, a tour of the facilities which the Review Panel "undertook", as also found in CS2.

However, the information provided about submissions in the ES is partial, sometimes ambiguous and some details have been redacted, as mentioned. For example, the exact number of embassies that "were invited to make a comment" is provided (13), and from what countries these were but the names or exact number of the "wide range of cultural organizations with which the School has relationships" are not included. Similarly, stakeholder groups making submissions are referred as the "School" group, the "academic staff" group and the student group. However, aren't academic staff part of the School, too? Or does "the School" include only people in administrative positions? Or could it be that "academic staff" refers to academics from other schools or other areas within the University?

Regarding meetings, while it is suggested that the submissions were received via invitation, there is no evidence about how those participating in meetings with the RP were approached as is the case in CS2 and CS4.

Nevertheless, a potential further submission may be discovered from the ES which at first sight does not seem so obvious due to the ambiguous way in which it is presented. When the writer comments that "Reviews of Schools *also* consider [...]"[emphasis mine] and

displays the areas on which the reviews of schools at this university focus (e.g. governance, resources, staffing, linkages, etc.), the “also” does not seem to have a clear grammatical reference in the text since the previous statement concerns not ‘other areas’ but the “series of reviews of academic programs and Schools” that the College “is conducting”. Therefore, what may “also” refer to? Could it possibly indicate that the “academic reviews” are a data source themselves? If so, an academic review of the School is neither attached to this report nor does the mention of such a document appear in the “Submissions Received” section. However, the following quotes in the Results chapter stating that “Bearing in mind the School’s stated belief that [...]” and “The panel noted the information provided in the submission made by the School XX [...]” suggest that such a document may in fact have been produced. Moreover, the following quote by (I) further confirms the existence of a School submission⁷⁰, and, in turn, implies a top-down approach to the review:

14. [...] and I was involved as the— as one of the 2 staff representatives in the sense of there was, in the panel, there was room for feedback from staff and I was appointed as one of the 2 people who would collect this information. So I had to talk to er more or less half of the colleagues in the school and er collect their thoughts, put it in a way in which er it was, I mean, it was to be anonymous and then present this to the review. Er I choose, because I am a social scientist, I choose to organize this in such a way that it was going to be feedback on a lot of levels.

The “Submissions Received”⁷¹ section in the report, while perhaps intended to complement methodological information from the ES, focuses on quantity and identification of individuals, which leads us to understand that the underlying purpose of this section is to demonstrate compliance with ToR and/or policy regulation. Furthermore, the information presented in this section shows misleading, inconsistent, vague and even repetitive information. For instance, while the title “Submissions Received” suggests that this section involves only submissions, information on meetings is also provided, and while the title may lead us to think that a summary of what is found in the submissions is presented, only a list

of names of the individuals categorized into three groups ((1) Staff; (2) “Student Submissions”; and (3) “Community or Group Submissions”) is provided. Moreover, submissions are only shown partially since those invited to make a submission are not mentioned and therefore a comparison of the response rate to such an invitation cannot be done. However, the methodological information is further restricted by the fact that the names and roles of individuals are redacted whereas the School or area of the University to which individuals are affiliated is not. However, other details are also redacted such as the brief comment next to the number of embassies contributing a submission.

Moreover, while names of staff are included, no names or any other details are provided about the students apart from stating the total number of submissions received (5). Lastly, it is noteworthy how in the Community category it is indicated that the “verbal submission” is followed up by a “written submission” whereas this distinction is not clarified in the rest of the submissions.

As for the meetings, unlike in CS2 and CS4, a schedule of meetings is not provided. Instead, superficial information, much like in the submissions section, is found where three (3) similar divisions evidence the staff, the students and the community members participating in meetings with the panel. From the staff⁷² group it is learned that two meeting formats (“group” and “individually”) were used and only the names of those participating in the individual meetings are identified whereas those in group-meetings are classified as “the staff of the School of XX”. Thus, a total of four individual staff members participated though, as specified by the writer, one of them did “(twice)” without adding any justification and their affiliation reveals an individual from a specific related College; one with a Graduate Studies role; one affiliated to the School under review and another School though to what College it belongs is not specified; and the one individual participating twice

is identified as a staff member of the School. Thus, the number of participants in individual meetings shows a considerably participation rate although their affiliation indicates a balance of internal and external participants, though the name of the group meeting reveals mostly internal people from the School.

As for the meetings with students, the only information provided is the total number of students meeting with the panel (12), with no reference about whether they were in groups or individual interviews or about how many were undergraduate, graduate, etc. as the ES points out. Additionally, the purpose of the information session with students mentioned in the ES is repeated here, “to explain the purpose of the Review, and to receive their [students] comments” as well as the name of the person conducting the session which has been hidden too. But, the question is would this reviewer, since this event was “held” by a member of the panel according to the ES, have organized this session which took place “prior to the review” in conjunction with the other reviewers? Was this member of the panel internal or external?

The final part of this section shows evidence about the meeting with “community” members though all that is indicated is the name of an individual(s), which has also been covered.

In sum, from the metaevaluation findings on evaluation methodologies in CS5 it can be concluded that first while a variety of audiences participate in this process as informants, there is clearly an interest in collecting the perspectives of staff over students. Second, that the purpose of the methodological information, rather than revealing rigorous methodological procedures, seems to seek to demonstrate that a variety of audiences participated in a variety of formats and to show the names of staff and external stakeholders participating in the review as stated in the ToR and perhaps as required by the

internal review policy though this is not indicated explicitly in the report. Furthermore, the necessary question to ask is the reasons motivating the decision to redact people's names since evidence of actual contents in submissions or discussions in meetings is minimal and in the Results chapter no name of participant is indicated when a reference to these data is made. What really matters in methodology is not the names of those who participate since for ethical reasons they are to be kept anonymous, therefore the need to do so may respond to policy requirements which in turn denotes a policy approach to the review.

CS3, as pointed out above, is itself a data source for CS2 though this is not reflected anywhere in the report and neither is any commentary made about methodologies used. However, some methodological information may be inferred from the Results chapter and Appendices, which suggests that data sources in this review consist solely of existing School documentation. This is clearly indicated by the names in the Appendices in the ToC⁷³ which correspond almost identically to the headings inside the report, thus revealing that their contents provide the primary sources on which findings are based. Examples of the variety of these documents include a summary of a School Board Meeting, various descriptions of governance structures, of the School's departmental roles, the School's current year administration plan, "financial performance" of the School including FTE (Full Time Equivalent) and SSR (Student/Staff Ratios) and two types of departmental review reports, namely, teaching quality reviews and research performance reviews. The only area inside the report not mentioned in the Appendices is "Relationship with external organizations and stakeholders" though this is easily excused since a list of external stakeholders is provided under this section.

However, additional existing documentation seems to have been consulted, particularly, University documentation as evidenced in the description the writer provides

about the University Economic Model (UEM) which was implemented as part of the University's Strategic Plan 2011-2015 to provide the framework for resource allocation throughout the University.

Moreover, anecdotal verbal data seems to also have been used as reflected in the quotes provided revealing the use of Blackboard sites. Regarding the perspectives of other stakeholders, the student responses from course evaluations are referenced in this section too but no additional perspectives from anyone seem to have been collected for the purpose of this review.

No further methodological information can be found or inferred from this report except for the evaluation model which is a "Self-review" as indicated in the title. However, the lack of evidence on how data sources were obtained or who was involved in collecting data and writing the report, the anonymity and the overall impersonal mode in the report makes it rather difficult to claim a participatory/collaborative self-review approach. On the contrary, as pointed out above, metaevaluation findings in this chapter on CS3 suggest the compilation and study of existing documentation and writing of the report by a single individual on behalf of the School.

In CS6, data sources and collection methods represent the main methodological information found. Apart from the "income sources" "modelling costs and benchmarking" stated in the ES, and the "extra data and information" in the Appendices mentioned in the Introduction, further evidence is found scattered across the Results chapter although some inference is necessary. For example, to address evaluation question 1, it must be inferred from the evidence provided about past reviews that the reviewers had access to the review reports. As an example, the writer shows in the following quote a comparison of the conclusions reached in two reports. Some conclusions of this review were radically different

from the 2007 review, despite the short elapsed time,⁷⁴ as well as by the explicit mentioning of the non-access to the recommendations of one of the reviews since “the [second language] language area was not involved in this review”.

Other data sources made explicit in CS6 ensue from the reviewers’ own past research. For example, the writer states that “We conducted an investigation into non-continuing students in [L6] in early 2010, prior to this recommendation” or regarding the proportion of the second language learners in schools showing “a huge growth in [L6] language learners at primary and high schools in recent years” which comes from the “Department of Education and Training” as indicated in a footnote.

Furthermore, the “extra data” used is illustrated in both the Results chapter and the Appendices. Thus, besides showing the actions taken to tackle recommendations from past reviews, additional data is used to contextualize the “actions and responses leading from each recommendation”. For example, for the recommendation to “urgently” change delivery to “2 times per week”, the self-reviewers confirm that the recommendation “was implemented” justifying it with a “detailed analysis of these changes and their impact” in a further section. Thus, sources of information in this example are also based on staff’s own work.

In the twelve Appendices, two kinds of extra data are presented. On the one hand, there is evidence supporting findings (10) and evidence of internal and external support for the program (2). Evidence of supporting findings are the result in some cases of the staff’s own work (e.g. Appendix 8.7 showing the dominance of the program’s exchange program in the region with two tables revealing eight exchange partner universities in comparison to much lower numbers in other programs). On the other hand, perceptions are collected from stakeholder groups in support of the program (e.g. Appendix 8.9 from the high school

principal with which the program has worked in a number of ways; and, as indicated, “48 pages” of letters by students both current and past supporting the program)⁷⁵.

Furthermore, besides the various limitations of the primary review from which CS6 ensues, in terms of the composition of the RP, the biased initiation of the review and the review questions, etc. commented above, other important limitations can be identified in connection with data collection for CS6. For example, in the “teaching costs vs income issue” section, which is key since “The review is firmly framed around finances”, the self-reviewers explain the difficulty to obtain data but which was finally made available “after repeated requests” though only four (4) days prior to the submission deadline of this report. A further limitation seems to concern errors in the data that were provided by the University which the reviewers feel compelled to correct (“We will here provide corrected figures, present details of missing information, and present more accurate revenues and costs”). (A) contributes evidence highlighting, on the one hand, the apparent initial inexistence of financial data on which the “allegations and assumptions” in the Review

Notice may be founded

A274. [...] so when I was told that, I went straight away to the business manager of the faculty asking for financial information of the (L6) program. She said ‘oh, we don’t have such things. We have only the degree and the faculty level but we don’t have the (L6) program, or languages’ and then what the hell the dean is talking about? He was talking from nothing.

which makes (A) perceive these as “created lying”. These findings are also supported by (C):

C7. [...] In the case of er of (U4) er for example, [...] we discovered quite quickly that the figures on which they were basing er their er their decisions at the level of the dean, the figures were inaccurate er very, very considerably inaccurate.

C75. [...] and they had bad figures⁷⁶

On the other hand, the difficulties in accessing data, to the extent of having to resort to threatening in order to obtain it, that is, making a legitimate demand look like a backdoor affair, and the rushed last-minute assembling of data is also emphasized by (A):

A284. [...] so I demanded financial information and they said ‘oh, it’s, it’s not (xxx), it isn’t sufficient, and it’s—’ I said ‘no, you said that ongoing deficit and a large amount of deficit. I need information’ and they refused to provide and they gave me the information, financial information only few days before the deadline of this response.

Q285. Are you saying that they did have information?

A286. Actually, they didn’t have it but so they have to make it but they didn’t want to give me because they were wrong.

Q287. After you requested it, they said ‘we don’t have it’ but after you requested it, they thought we better get something together, put something together here.

A288. Well, they didn’t, so I went straight up to the registrar of the university and ‘this is outrageous, and if they claim that and if you conduct a review based on the financial problem without giving me financial information, how on earth am I supposed to— they will really compromise university’s credibility. (xxx) I’m gonna go on the media’. And then he— the registrar acted on that. Er I got information just 2 days before, but I—

A290. [...] by the time they gave me information I had all analysis done and ready to roll but they didn’t give me any information, I had to go backway, and for the revenue I had to go to the university statistics office to get data because—

A292. Uh, no, that statistics department understood what’s happening and they said that that’s outrageous so they quietly gave me the information.

Q293. Because the financial office refused.

A294. Yeah, and the financial office through faculty gave me the information 2 days before the deadline.

Regarding the evaluation approach, it may be claimed that on the one hand, evidence about the primary review reveals a top-down approach to the review, whereas the various references across CS6 in the first-person plural associated with data analysis such as “our own analysis”, “our own research on course structures”, “we conducted a survey amongst our students”, “Now we benchmark our student enrolments” etc. clearly indicate a self-review approach. However, there is no evidence in the report about the extent to which each of the two self-reviewers were involved in what parts of the conduct of the review,

analysis or writing of the report, although, from (A)'s verbal data below it may be inferred that (A) is behind most of the data collection, analysis and even the writing of the report:

Q247. Did you address [all of these questions?

A248. Yeah,) probably. Well, I wrote 84 pages document for that ((laugh)) which they didn't expect. They gave me 2 pages [...] so this is executive summary.

A266. Sure]. Yeah. This one, I made a 2 version. Everything I wrote is my work and anything, other than financial information—

Q267. You were the convener of the program at that time.

A268. Yeah, I wrote the whole thing so um there are a few things like er— [...] particularly the fact that (A)'s interest in updating the new Dean ensued more than anything else from suspicion due to past experiences with university management malpractice and lack of respect for staff.

In CS7, a general methodological description is found in the Results chapter focusing mainly on data sources as well as data collection methods, although some scattered comments indicating limitations are also present. The description explains that findings are demonstrated “based on an analysis of” a variety of data sources either existing or collected via different methods and stakeholder groups (“program data, feedback from students, ongoing and from focus group discussions, and from external stakeholders”). However, the use of a further data source is suggested by the following statement: “While anecdotal evidence suggests [...]”. Among program data can be found benchmarking of EFTSL⁷⁷ enrolments, CEQ⁷⁸, range of languages, TERs⁷⁹ for language degrees, and promotion activities. Moreover, descriptions of other language programs of local, regional and national universities are also provided thus indicating further data sources. Finally, “mapping analysis” is conducted “within and across all courses focusing on course teaching and learning arrangements, objectives, content and assessment” in order to study program coherence. Evidence of these data to support findings is used in the Results chapter rather than in the Appendices, as found in CS3, and it consists mostly of tables containing numerical data.

Furthermore, commentaries about methodological limitations can be found regarding data sources. For example, when discussing Tertiary Entrance Rank (TER) cut off scores (used to rank students for entrance into competitive university courses), the data available for a particular period of time is considered “less reliable” since some students “are not recorded as having a TER” which makes it “problematic” for analyzing whether “students are high achieving or not”. When considering the area on “Student profile, equity and inclusivity”, a finding is mentioned for which “data that would suggest an explanation for this” “we do not have”. Other explanations are offered also regarding, for example, problematic comparisons of scores across programs at other universities “because of the way data is reported”, thus justifying that, instead, scores are compared “with data from the most relevant FoE⁸⁰ code”. A further example revealing methodological limitations is when discussing “Graduate satisfaction and GCEQ⁸¹ performance” which appear “very high for the two years for which data is available” but which is based on data provided by “relatively few students” [...] which is then justified by alluding to the young age of the program.

(R) highlights a broad limitation, which occurs “often”, in connection with data collection which involves the difficulty to access data from other universities and justifies it stating that “it is after all a competitive system”. This finding raises the question about whether competitiveness may be the reason that justifies the general reluctance by University authorities to grant access to review reports for the purpose of this research.

Regarding the evaluation approach though not discussed, earlier metaevaluation findings and verbal data stressing, for example, that review activities were “undertaken by the teaching team” indicate a participatory self-review approach. However, while the term “team” is mentioned throughout the report under different denominations (“the [program name] team”, “the [program name] teaching team”) and even a number of its components

are mentioned (“the fourteen or so staff in the [program name] teaching team”), the members of this team are not identified. However, the quote “Integrated with their languages studies, students also complete courses in applied linguistics” in the first section of the Results chapter denotes that the “team” may be composed of staff from both the languages and applied linguistics areas.

Interestingly, according to verbal data, the collaborative approach emanates from institutional policy, thus suggesting compliance with review policy

R8. [...] in my university, every three to five years, a program must be evaluated formally by its teaching team, so, all faculty contribute.

R10. [...] There is a policy, our policy is (policy reference). [...] The teaching team, all of those involved, are part of the evaluation.

and it seems to be circumvented at times, though no explanation is provided as to why:

R16. [...] in some programs, the program director just sits down, looks at the data and writes it.

Moreover, while commenting on the collaborative approach of CS7, (R) feels the need to justify how this review “which is conducted totally internally” is not so much so but “well, largely internally” because “there is an external element” which signals (R)’s awareness of the pitfalls of using an internal-only approach to evaluation. But what does this “external element” involve? According to (R), it consists of the various committees of internal members of the university at different levels who check and endorse the report prepared by the program (“So there are three layers, there are points of accountability”). However, can these internal committees be considered “external”? The following quote discussing the ‘public’/‘private’ nature of the evaluation report shows that for (R) they are:

R12. [...] That report that you have seen is not a public document, but nor is it private and internal because of the three accountability check points: the school, the division or faculty and university wide. So, the eyes of other members of the university community who had nothing to do with our languages program were going to be looking.

3.6. The Results Chapter

3.6.1. Description

In an evaluation report, this chapter represents the central part of the document, presenting a descriptive account of the findings from the investigation carried out by the evaluators. Thus, the ERC asks whether findings have been described clearly and logically; whether tables, charts and graphs are understandable and appropriately and consistently labelled; whether the discussion of findings is objective and includes both negative and positive findings; whether all evaluation questions are addressed or an explanation is included for those which are not; and whether findings are adequately justified.

3.6.2. Findings

First of all, it must be noted that the similar references in the third person in CS2, CS4 and CS5's Results chapter such as the Review Panel "considered", what "concerned" them, what they "agreed", what they were "(very) impressed by" what they "saw a need for", what they "felt" and what they "commended" and "recommended" suggest that the information under this section is not necessarily the reviewers' findings but a summary of what the reviewers found as written up by the same anonymous author of the previous sections. The fact that the third person used in the ToR is found throughout the report in these three case studies, in other words, that there is no distinction between authorship of the ToR and the rest of the report implies that the writing of the report may be associated with an anonymous administrative internal member of the University (thus = potential authorship in CS2/CS4, CS5). While this important metaevaluation finding raises questions

which at this point cannot be answered (e.g. the reason why the reviewers' report is replaced by a summary, or who the recipient of this summary is, or why the original reviewers' report is not made available for the purpose of this thesis).

Moreover, given the deep level of analysis required in this chapter, the metaevaluation findings from the rest of the reports, that is, of CS1, CS3, CS6 and CS7 are hereby dealt with case by case starting with CS1 and CS6 since these are the ones that offer the best composed and most logical account of findings.

3.6.2.1. CS1

In CS1, all the questions are addressed, and tables and figures are used to illustrate findings. However, while all tables are discussed in text, not all of them are labelled.

To clarify the structure of the discussion, the analytical questions are restated under each section together with a brief introductory explanation which helps contextualize the question. For instance, before the discussion on the reasons why discontinuation may occur, the reviewers emphasize taking into consideration "the characteristics of this singular student cohort" whereby the program is viewed as "expendable" since it is offered as a supplement and not as a degree.

A general logical pattern can be identified in the presentation of findings for every issue that is addressed. For example, in the findings from the qualitative data section, the writer provides an overall statement about perceptions about the program which convey positive findings on the part of students whereas positive and negative findings are revealed through staff' perspectives⁸². This discussion is appropriately supported with objective evidence such as literal quotations from participants' verbal data. Sometimes this evidence is quoted in-text whereas other times it is separated from the main text depending on the

length of the quotes. Moreover, every time a quotation is used, reviewers include systematically information on the participant. For example, for students, the language, the entry level, their continuing/ discontinued status, and years of study are noted. For staff, the label “staff” and the language they are associated with are also placed next to the quotation.

When contradictory findings are encountered, the reviewers dig deeper into the data to find out whether they can be justified. For example, the finding revealing positive comments about teaching methodologies but also the finding that there are “inconsistencies” is justified with two extra findings from the verbal data, namely, (1) the “poorer teaching” occurring more at higher levels with instructors who either manifest characteristics that do not lead to quality teaching or who use inappropriate “teaching devices” (e.g. interaction limited to Q/A, teacher centered rather than student centered); and (2) the “disparity of views” on a common communicative teaching approach (e.g. some staff supporting “communicative teaching emphatically” while others viewing it as “too reductivist” or as a method which does not encompass the goal of language study, that is, to engage with “cultural studies” or which does not contribute to solve the “severe problems with accuracy”).

Sometimes, the reviewers resort to research to justify findings, though without referencing, as mentioned, and when appropriate, professional advice is provided. For instance, regarding the staff’s perspectives on the communicative approach not providing enough attention to accuracy, the reviewers recommend a method which allows the development of communicative skills at the same time that it focuses on form (“Focus on Form approach”). At other times, a specific view on particular issues is also manifested. For example, they show disagreement on the perspective that the only purpose of language

studies is to engage in cultural studies and tie students' insufficient command of the language for cultural studies to a broader issue. However, in "Timetables" which emerges in the student data as "a scourge" and a problem due to complaints about teachers not being flexible enough, a particular inclination can be identified in the reviewers' judgment. For example, in distinguishing "enlightened" teachers who understand that languages are not a degree subject from other "less helpful" teachers signals a tendency to narrow the issue down to teachers' response to timetabling difficulties and not consider, for example, revisiting the current timetabling model or at least raising the possibility of an inquiry into a different one, asking questions such as 'could other timetabling models to facilitate language study arrangements around campus be looked into?'

3.6.2.2. CS6

In CS6, both the evaluation questions from the Review Notice as well as the tasks that the self-reviewers announce in the Introduction are all tackled in the Results chapter. Moreover, tables and charts are often used to illustrate findings, which are all discussed in the text and adequately labelled.

Similar to CS1, the questions and a brief commentary introduce each section to either explain further the structure and contents of the discussion of findings:

We will first re-state the nature of each past review and important features of them, and then present our actions and other responses leading from each recommendation. We will then separately address the four specific points in the question above.

or to provide a context for the question, as in question 3, where the question

“What other steps could be taken to eliminate in future years the serious deficit incurred over a number of years in running these units?” is situated within the conclusion ensuing from the previous chapter highlighting the erroneous assumption in the question:

We have clearly demonstrated that there is no deficit. We have brought a net income to the Faculty. We have slight concerns that the cash surplus we generated was lower in 2011; but this is partly due to the fact that (P33)’s⁸³ workload has been increased from 1.5 to 1.6. Also, we have demonstrated that the net financial effect of the Japanese language area on the University as a whole, and on the Faculty, is very positive. [...]

Therefore, while the RN’s evaluation questions lead to focus on negative findings such as “low enrolments”, “deficit”, “high attrition”, etc. these are difficult to justify since, as the quote above explains and as stated in question 1 section about actions taken to implement recommendations of past reviews, “The financial viability of [L6]⁸⁴ Language offerings was neither the trigger nor the focus of any of these reviews”, whereas in CS1, evidence shows a high discontinuation rate. To illustrate this, the self-reviewers describe in detail every action taken and research conducted to justify findings and in turn to emphasize compliance with University expectations. For example, the discussion for every past review is, in every case, structured in the same fashion. First, the reviews are identified separately by naming the year in which it was conducted, and second, an introductory note presents the name of the review, the name of the reviewer or individual who “led” the review and the number of recommendations that involved the program. After that, the recommendations are analysed individually, each containing a heading which represents the focus area of the recommendation. Under these headings, the self-reviewers reproduce literally the recommendation and an explanation about the extent to which it has been implemented and why, how it has been implemented and whether further work needs to be pursued.

To prove further compliance with institutional priorities such as exchange programs as inferred from the quote below, the self-reviewers study the numbers of inbound and outbound students at the University. These are presented in the Appendices and are intended to complement the pie chart and explanation in the Results chapter which lead the writer to conclude that:

Our contributions to the exchange program are extremely significant in the context of the [U4]⁸⁵ policy on International Educational and the Strategic plan.

Therefore, it can be claimed that positive findings ensue from proving wrong the negative assumptions in the questions. For example, in question 1, the reviewers show evidence of how program staff not only had acted upon recommendations but also even done work prior to a recommendation which, in some cases could only be partially implemented due to University-wide degree constraints. Negative findings though, by implication, are associated with actions by the initiators of the primary evaluation as shown above (e.g. contradictory conclusions from one review to another; errors in the reports). In fact, this is a pattern reflected throughout the Results chapter. For example, in the “Attrition Issue”, after communicating what is to be found under this section

We detail below various efforts made to gather data, analyse the situation, and implement change to this end over a number of years

actual evidence that justifies findings is provided in the form of numerical data from past investigations about “true retention rate for the major” and from first year student numbers over eight (8) years. This evidence is displayed clearly in various tables, complemented by other tables in Appendix 3. Moreover, actions taken in order to retain students, particularly in the first year since “we have no control and little influence over degree structures”, are also described in fine detail, as well as the application of eight (8) different strategies to encourage retention. Examples of these are “We established a peer mentoring program

[...]”, “we performed a major restructuring of the contents covered [...]”, etc. or strategies such as “increasing collaboration” with a specific Faculty that in the future may influence enrolments positively, thus showing how the program anticipates actions in order to retain students despite structural constraints.

In sum, positive findings reveal compliance and beyond and the negative ones reveal unprofessional practices and/or lack of information from those managing the university and commissioning reviews.

3.6.2.3. CS3

In CS3, the Results chapter represents 99% of the whole report apart from an incomplete ToC and the Appendices, the existence of which is known from the ToC but which have not been made available for the purpose of this research. The areas under review in CS3, as Table 6 below shows, replicate to a large extent those in CS2’s ToR. While according to the table two areas (2 and 9) in CS2 appear not to have been addressed in CS3, a careful analysis reveals that in fact, the internal organization of departments and programs is tackled in the Introduction. Regarding CS2’s area 9 ‘Internal and external communication and marketing practices’, no independent section is dedicated to this particular topic nor any explanation provided. However, from the persistent emphasis on the accelerated language program throughout the report and the contents of the external relationships discussion it may be inferred that the writer feels that the School has in fact complied with external communication and marketing practices.

CS2	CS3
1. Governance structure	1. Introduction

2. Internal organization into departments and programs	2. Governance structure
3. Professional support:	3. Professional support
4. Financial performance and management	4. Financial management and performance
5. Management of curricula and teaching	5. Management of curriculum and teaching
6. Quality of research support and research performance	6. Research support and performance
8. Relationships with relevant external organizations and stakeholders	7. Relationship with external organizations and stakeholders
9. Internal and external communication and marketing practices: effectiveness	8. Conclusion

Table 5 Areas under review, CS2 and CS3

From the above metaevaluation findings, then, a certain lack of clarity in the structure of areas may be inferred which can be found in fact in other parts of the chapter. As an example, in “Management of Curriculum and Teaching” while the “Diploma of Languages (accelerated mode)” is placed under “Innovation in Teaching and Learning Activities” as an illustration of the School’s efforts in this area, the reasoning behind not including other similar activities such as “Teaching Day”, “e-Learning Incubator” and instead placing them at the same level as the rest of the subsections is unclear, unless the writer intended to give these areas a particular emphasis. Moreover, the “In-country study at [2 foreign universities]” appears as an illustration under “Innovation in Teaching and Learning Activities” when it could have been easily situated under “Exchange and In-Country Study”. Table 7 helps to visualize these findings:

SECTION 5: MANAGEMENT OF CURRICULUM AND TEACHING
5.1 DEPARTMENTAL CURRICULA REVIEW
5.2 FIRST YEAR EXPERIENCE AND STUDENT SATISFACTION
5.3 LANGUAGE PEDAGOGY
5.4 TEACHING DAY
5.5 eLEARNING INCUBATOR
5.6 CONTRIBUTION TO PG COURSEWORK PROGRAMS
5.7 JOINT TEACHING OF PROGRAMS ACROSS THE FACULTY AND OTHER TERTIARY INSTITUTIONS
5.8 EXCHANGE AND IN-COUNTRY STUDY
5.9 BA LANGUAGES
5.10 INNOVATION IN TEACHING AND LEARNING ACTIVITIES
DIPLOMA OF LANGUAGES (Accelerated Mode)
IN-COUNTRY STUDY AT [2 foreign] UNIVERSITIES

Table 7 Structure of Report, CS3

Other times, the structure of the report is not consistent. For example, some of the sections have a brief introduction but not all of them do, and while two of the sections finish with challenges as a kind of summary/conclusion, the rest of them do not present any concluding remarks.

More importantly, the lack of explanation also affects a full understanding of what is being read. For example, when the “Five principal strategic objectives” of the School (e.g. Research, Service, Curriculum Development, Language Teaching and Technology) are defined in the Introduction, the only comment made about them (“strong progress has been made in implementation of the operational plan for each area”) clearly reveals an interest in emphasizing compliance but it is unclear whether the findings that follow are presented as evidence that justify this judgment or not.

Furthermore, given the similarity of CS2 and CS3's sets of areas, it could be argued that these may be part of a specific University Strategic Plan mentioned in both CS2 and CS3, though no evidence is offered in this regard. However, the findings below do show a general underlying purpose in the description of findings to prove compliance, which explains why it is rather difficult to qualify CS3's findings as negative or positive. For example, can the descriptions provided under "Governance Structure" of the School's several governing bodies such as the "School Board", "Executive Committee", etc. be considered findings? Likewise, under "Management of Curriculum and Teaching", the writer fundamentally provides rather than a discussion of positive and negative findings, a mere description of ten (10) specific areas demonstrating what the School has done in those areas.

Yet a few positive and negative findings may be discerned in some parts, though still presented as illustrations of compliance. For example, in "Departmental Curricula Reviews" the writer emphasizes, first, that the School is in fact conducting these reviews, and, second, the implications from the Review Panel's recommendations which the writer presents as positive (e.g. "practically all Departments and Programs ended up with leaner⁸⁶ and more streamlined curricula"). A closer look at these recommendations reveals that these involved (a) a Formal Review of some language areas; (b) internal restructure in some cases; and (c) cutting low enrolment units or units "with no strategic relevance for the curriculum" which is believed to occur "in numerous cases". Therefore, it could be argued that the positive finding reveals compliance with regulated reviews and efficiency in accordance to the UEM. But, how is a *Formal Review* defined? Must it be assumed that the curricular reviews⁸⁷ were 'informal reviews'? As a matter of fact, a Formal Review seems to be associated with a predetermined change: the department to which this Formal Review was recommended

had had its former area of expertise “changed”. Therefore, it may be claimed that reviews, whether formal or informal, are used to make changes to align evaluands with University priorities.

Other times, as in the following example, a finding is presented as positive and negative at the same time and in relation to efficiency. For example, the “successful centralization of the administration structure of the School” is described in terms of its positive effect, that is, “a relatively low turn around”, whereas at the same time the reduction of resources, that is, “administrative support tasks [...] devolved to School level with no provision of additional resources” is also pointed out.

The justification of findings in CS3 is sometimes vague and resembles the summary character of the findings section in CS2, CS4 and CS5. For example, the results of student course evaluations are presented with a conclusive comment stating “the very positive” evaluations of students which included commendations to teachers with no further comment. Nevertheless, once more, in this section, the focus seems to be the demonstration of compliance. For example, the writer begins by highlighting the commitment of the School to “improving the first year experience” and its active participation “with that aim” “in all University and Faculty led initiatives” which is supported by the praising comments of a Dean as mentioned (e.g. blended learning and Blackboard; the reporting of results on student course evaluations). Therefore, compliance is associated in this section with ongoing work towards attaining positive student evaluations, prompt reporting on student course evaluations, high student response rate, and involvement in technological initiatives.

As another illustration of weak justification, under “Effectiveness of the Interdisciplinary Programs”, two main programs are compared. The effectiveness of the first

program is measured by stating that it is “one of School’s major successes” whereby success is associated with range of units, and range of participating departments, and with enrolments. As for the second program, it seems that it is not successful since what is described is the length of the program, its restructure, the role of a “Professor” “the School has whose task is to redesign the program”⁸⁸, the date of completion of such task, the “two junior units” ready to be introduced and the “Senior Intermediate units” developed to enable students “to follow pathways” in other disciplines. Interestingly, the concept of success is clarified in the following section “Challenges Programs and Departments Face” where the three reasons why the second program is not successful, namely, the traditional practice of offering “many of their cultural units in the target language”, not counting undergraduate units towards the major, and the competition with another department, clearly underlie the issue of enrolments. Therefore, the challenge, that is, what compliance means in this case, is to model the first program to increase enrolments since “The School’s intention is to develop” a program where students are not taught in the target language but rather are “encouraged to learn the languages of the region”.

In the last area regarding external relationships, the findings are superficially justified and what stands out is the orientation in this section to demonstrate efforts made to establish collaborations on the one hand, and marketing practices for the purpose of acquiring funds on the other. For example, the “impressive record of collaborations with external organizations” is justified with a cursory description revealing the establishment of a recent “collaboration” with an outside university, with a “successful partnership” with another foreign university, with “the very close and fruitful collaborations” with a particular foreign language association, etc. Similarly, the “very active and visible” role of the School in

“external engagement” is stated where “community engagement” is associated with a list of providers who offer “very generous donations”.

Similarly, while CS3’s Appendices seem to have the purpose of providing evidence supporting findings which may explain the absence of tables, etc.⁸⁹, the sources of findings and the relationship of findings with evidence in the Appendices is not clearly indicated. For example, findings from “First Year experience and Student Satisfaction” may be assumed to be based on data from the Departmental Quality of Teaching Reviews in one of the Appendices but this is not indicated. Likewise, it has to be assumed that the data source for percentages of student response rate is an institutional database as suggested by the following quote:

For example, in semester 2, 2013, the response rate was above 70%. In comparison, one of the FASS Schools had a response rate below 50%, with the other Schools sitting between 53% and 61%.

3.6.2.4. CS7

In this case, as in CS3, the Results chapter constitutes almost the entirety of the report and in light of the institutional policy driving this review according to (R)’s data, the areas under review, it must be assumed, ensue from that policy which embodies domains such as the “strategic considerations” explained above but also “Demand and market position”; “Program viability”; “Retention and success”; “Satisfaction of undergraduate students, graduate students and external stakeholders”; and “Academic design”.

In contrast with CS3, the structure of findings in CS7 is more consistent even though subsection headings come unnumbered. For example, the subsections that branch off from every main section finalize systematically with “Proposed Actions” (PA) and findings are

logically discussed. As an illustration, after general descriptions of programs from other universities both from the same state and from around Australia are presented, a final description summarizes the “Features of similar programs” from which “Issues” and “Other threats” are extracted as the following quote announces:

Issues facing [program name] arising from an analysis of competitors and similar programs, therefore, include: [...]

In turn, these findings provide the basis for the PAs that follow.

In CS7’s, as in CS3, positive and negative findings are described in terms of the extent to which the program is aligned or needs to be aligned with priorities although in some examples this is more obvious than in others. As an example, in “Academic Design”, a pattern can be identified in each of the three subareas presented, namely, (a) “Program coherence”; (b) “Academic design”; and (c) “Program support”, in which an emphasis is clearly placed on how the program takes the institutional framework seriously and orients the program towards meeting the goals established by this framework. For example, in “Content, workload and assessment”, it is stated that the program’s design “supports the development and integrations of research experience, and skills, experiential learning and academic literacy” according to the “Teaching and Learning Framework”. In “Academic Design”, the same pattern gets revealed. In fact, the very first sentence in each of the subareas under this heading starts with the assertion that the program meets expectations (“The graduate qualities are well integrated”; “The teaching/research nexus is a primary focus of the [program name] program”; “From 2008, the applied linguistic major began to introduce on-line courses”; and “Strong focus on indigenous content in key courses”). It must be noted that in all these areas, findings are described and adequately justified. For example, as evidence of the focus on “the academic literacy approach to develop each

student as a critical consumer”, a table is provided with the “course” which has integrated this approach, in this case four (4); a description of their different “focus”; and the “timing” within the course sequence.

In other places, however, compliance with priorities needs to be inferred. For example, the results from “The Graduate Destination Survey” reveals an interpretation of findings whereby a need to “address” employment of students is perceived as necessary. This need is further shown in the five PAs where PA 2, 3 and 4 involve working closely with various state organizations (“the Chamber of Commerce” and “the Australian Public Service Commission”) to establish a number of activities that will contribute to better match the program with employment opportunities. And PA 1 and PA 5 show activities that the program intends to take up itself to also contribute to that same purpose which involve following up on alumni and incorporating a “Career Management” component into the program. Therefore, it is clear that the destination of graduates seems to be a University priority to implement, and therefore the program is demonstrating in this section efforts towards ensuring that this area is tackled.

In CS7, findings are generally adequately supported with evidence, and tables are used regularly to illustrate findings which are clearly presented, described in the text and labeled consistently. For example, in “Graduate Satisfaction and Outcomes”, findings from the first subsection show positive results from a student evaluation instrument, namely, Course Experience Questionnaires (CEQ) revealing “a high level of satisfaction with the program” justified with “Student comments from the 2009 CEQ”.

In “Demand and Market position”, findings are justified on the one hand with data from five (5) tables showing mostly numerical data. The first four show benchmarking results against “our competitors” that is, with two other local universities using four sets of

data, (e.g. CEQ (Course Evaluation Questionnaires) data; “Total language enrolments”; TERs; and “Languages available”). The fifth table, additionally, contains the program’s TER data showing number of students per cut off TER scores (in 10 categories), number of commencing students with TER, number of commencing students with TER unknown, and number of continuing or non-school leaver students in four different years, and finally, an “All enrolment Years” showing the total number of enrolled students for every category. On the other hand, another variable taken into consideration for benchmarking is “Promotion” where evidence of marketing plans, a past one⁹⁰ and a future one, for the program are explained in detail and which are in fact attached as Appendices to the report. Moreover, promotion is perceived by the writer as a strategy to “address the issue of student numbers and the TER cut-off”, but also as a need which comes from external demands such as the “regulated environment from 2012” as well as “advice provided by stakeholders”, from “Evaluation feedback from participants [...]”.

However, while “Retention and Success” could also have used numerical data as evidence, it does not, and the reasons for not including it are not explained. Instead, that retention exceeds the Divisional average in all but one year is presented as the main finding and presented as a concern given the evidence on positive student evaluations. The justification is shown via a student comment “in the focus group discussions” and “the stakeholder discussion” which highlights how students sometimes may “feel lost within the range of options offered [...]”. Interestingly, the action suggested in the discussion is to “surpass the division benchmarks for retention” which raises the question as the need to “exceed the divisional average” unless it is a mandated strategy.

In fact, the general theme of compliance is revealed through the very names of sections and subsections. For example, collaboration with other University programs is

inferred from the section “relationship to other programs in the University” where the description focuses on the potential connections with other areas and the current connections already in place providing specific examples such as ESL⁹¹ and a pathway to a Masters of Teaching which is supported by “stakeholders” who offered to promote it. Additionally, the theme of enrolments emerges throughout the report, in particular, strategies for attracting students. For example, besides “Market Demand” described above, “emergence of any new markets and how the program might exploit those markets” shares the same topic where the discussion distinguishes International and domestic markets and focuses on targeting students emphasizing the marketing plans designed in order to attract students. Also, there is a clear association of the viability of the program with enrolments and the need to increase enrolments. Thus, in “Program Viability” the topics that are tackled in the subsections, namely, “current student profile”, “access and participation of equity groups”, “further data relevant to the standing of the program (e.g. stakeholder views, age of program, strategic importance, relation to competitor programs)”, and “any trends that may indicate improvement” reveal an interest in demonstrating compliance with issues such as diversity of student population, increasing program competition for attracting students, and the pursuit of strategies that may help increase enrolments.

3.7. Summary, Conclusions, and Recommendations

3.7.1. Description

The ERC recommends that there should be a concluding chapter in which, as the name suggests, a summary of findings and a conclusion are included and, if appropriate, recommendations. However, according to the ERC, summaries may be found at the end of

each section within the Results chapter. What is important is that the summary includes a discussion and interpretation of findings as well as judgments about the program that cover merit and worth; that both the summary and the conclusion reflect findings fairly; and that the recommendations are based on findings.

3.7.1.1 Summary of Findings

Due to the fact that CS2, CS4, CS5 lack a Results chapter, it is not possible to study the extent to which their summaries and conclusions reflect findings fairly and whether recommendations are based on findings. For this reason, in this section of the metaevaluation report the findings from the summaries of these three reports are first discussed as a group to study potential shared traits.

Summary in CS2, CS4, CS5

The descriptive summary found in these three (3) reports is of similar length (8, 10 and 8 pages respectively) and is based on the various reviewed areas of the program as the separate headings in each of the chapters reveal.

In CS2, a pattern of meaning can be identified whereby the discussion focuses on merit and worth judgments emphasizing compliance with priorities (“research outcomes”, “innovation”, “cross-disciplinary”, “cross-listing”, etc.) associated with finances, (“income”, “viability”, “sustainability”). In other words, it consists mostly of a description with an interpretation underlying a purpose to demonstrate the extent to which departments within the School are viable and serve the University strategically, and also whether strategies have been applied to align them with University priorities, particularly to make them sustainable.

For example, in “Research” the writer starts by concluding that “It is clear that” strengthening the research culture has been a priority, as proved by the “significant gains in

research performance” (though “moderate publication counts” versus “significant growth in grant income” are also mentioned), as well as other “Particularly pleasing” examples (e.g. cross-disciplinary research activities, the degree of research mentoring, increasing support for inter-departmental research networks, participation in Faculty-wide research initiatives, and various research support activities). These, the writer insists, are “positive” examples of “growing recognition of the need to prioritize research outcomes”. In “Curriculum”, similarly, there is a clear purpose to demonstrate compliance where “diversity”, and “accountability” associated with viability and sustainability frequently emerge as themes. Hence, “diversity” (range of programs, diversity of programs’ “approach”, “strong suite of programs”, “comprehensiveness and diversity of programs”, “diversity as a defining characteristic of the School”) is presented as positive as well as a way to respond to challenges (“positive response to change”) associated with decreasing enrolments (“waxing and waning of student interest”). This capability, too, is praised as a characteristic of the School to demonstrate a willingness to account for the viability of the curriculum (“Collective acceptance”, “clear understanding at all levels of the School” of the “need” to balancing “suite of programs against issues of sustainability”, also defined as a balance between “academic integrity” “against operational and strategic imperatives”, or of “accountability for the viability of curriculum”) which is exemplified with the discontinuation of “non-viable Masters of coursework programs”.

A further illustrative example of the School’s efforts to make the curriculum viable is, particularly, the languages accelerated program which is contrasted with the Diploma of Languages. Thus, with the program, the School shows the recognition of “the opportunity” to widen the range of students, which implies an increase in enrolments, a “balance of innovation and strategic focus” associated with the “staged introduction of specific

languages”. So, it may be claimed that “strategic” in this case may mean to select those languages which attract a particular student profile and/or to attract the highest number of students. Nevertheless, the association of this program with finances is most clearly evidenced in the “note of caution” from the panel not to let this program become a “cash-cow” strategy. As an additional illustration, the description of the e-Learning Incubator⁹² demonstrates once more compliance, that is, “a commitment from within the School to encourage innovative thinking with regard to delivery of its programs”. The underlying idea of compliance is further highlighted when explaining the support of “The Teaching and Learning Committee” involving promotion of “innovation (blended learning) and educational technology.”

These positive examples are followed by a description of three (3) School-wide programs where the issue of viability in connection with finances emerges once more. For example, one is described in terms of being “small but well regarded by staff” with issues due to “uncertainties around access to staffing”; a second one of being “financially robust”, thus making “a substantial contribution to the School”; and a third one “yet to prove itself financially viable” respectively⁹³. However, other programs described underlie the same financial theme. For example, one more program’s discussion revolves around the issue of its “positioning and sustainability” and three further departments are brought up as concerns due to “sustainability of curricula and staffing”.

In CS4’s summary chapter, the summarizer consistently uses a pattern revealing not so much a discussion, but a descriptive summary of positive findings followed by negative findings which lead then to a specific recommendation. These descriptions reveal judgments of merit and worth, too. For example, in “The Diploma in Languages”, the description begins with the “very successful” status of the program followed by a description of its limitations,

and ends with two main recommendations. Similarly, in “Study Abroad and Exchange Programs”, the positives are first commented on before moving on to the issues and the recommendations that ensue from the issues.

Two main themes seem to emerge from these descriptions which are intrinsically interrelated. On the one hand, the justificatory, sometimes even defensive tone, and, on the other, the focus on marketing, student enrolments and finances. For example, in “The Diploma in Languages”, the positive findings reveal justification of its existence by stating its beneficial contributions, which focus on access by students from “all faculties”, “the development of a language culture”, what would happen to students wishing to study a language should this program not exist (“Without this program, [...]”) and the support to this program from “all faculties” who see it as an added “value” in terms of expanding the students’ “employment possibilities beyond Australia’s shores”. These findings reflect at least in part a concern with educational values. However, the description of negative findings centre around the “limitations” of the program revealing its unavailability to international students, lack of program coordinator and insufficient marketing by the Faculty. While these topics provide clear evidence of the focus on the financial aspect of the program through an emphasis on enrolments, it gets more clearly revealed when, for example, the summarizer adds the possibility of allowing access to past graduates according to what “Some staff believed” and also access to students from two other regional universities as “it was argued”. Furthermore, while the lack of a program coordinator is associated initially with a lack of “defined system for the provision of course advice and promotion of the program”, interestingly, the “cohesion” to which a program coordinator would contribute is further down perceived as a role for “marketing the Diploma” and oversight of areas such as retention rates, completion numbers, enrolments, and tracking of

both graduate employment outcomes and satisfaction. Therefore, cohesion is associated with administrative more than educational related issues with a focus on enrolments which, according to the writer, the Review Panel believes will result in “cross language engagement and increased dialogue among staff”. Lastly, after the list of the tasks that such a program coordinator would take up and the recommendation ensuing from it is presented, the section ends with a further comment revealing once more the focus on student numbers since it describes a further inhibition to access the program (e.g. a Faculty requirement demanding to complete a semester before enrolling) which also leads to recommending its removal.

A further illustration of the persistent focus on enrolments is the section “Study Abroad and Exchange Programs”. While this program is justified by its popularity (“appear very popular among language students”), the meaning of popularity is associated with “the value” students and staff perceive about this program but also with enrolments. For example, student percentages are mentioned to “vary” according to discipline but overall a “significant number of students” “will undertake” this program. In fact, the concern that is identified about this program is once more student access, due to funding problems for which the Faculty is recommended to search for external “possibilities for funding”.

In CS5, more than a summary of findings, it is fundamentally a summary of recommendations with findings being used as the basis to justify⁹⁴ recommendations. In other words, the discussion and interpretation seem to be done for the purpose of giving grounds for a recommendation, thus, reinforcing the underlying concept of compliance. For example, in “Range and Delivery”, the recommendation ensuing from this section suggesting the exploration of “the pedagogical strategy for modularising the constituent parts of language courses” ensues from the very last issue discussed in this section, namely,

background speakers' "differing level of competence". While no justification is provided, since the adoption of this strategy would affect not only heritage speakers but all students, the driver for modularization may be associated with a broader issue mentioned at the beginning of this section, that is, "retention of students", and therefore, as a way of providing a more flexible curriculum which would contribute to secure student numbers. In fact, this strategy may very well respond to this aim since it is connected with "admitting students at differing levels for each of these competencies". The flexibility of this type of strategy goes hand in hand with a further suggestion, namely, "online or other modes of flexible delivery" regarding the offering of tuition in two specific languages, without explaining why these languages and not others, since the "market for these languages is not likely to be large". Therefore, in the changes that are sought, flexibility, which emerges explicitly or implicitly in other parts of the report, too, while potentially an educational matter, is clearly associated with the issue of student numbers.

Interestingly, the nature of some of the recommendations previously presented reveal the issue of funding ("mechanisms for funding the period abroad") as well as compliance with University priorities. For example, the writer confirms that flexibility is to be introduced in line with" the University's "educational planning", and that the suggestion proposed to model another College's overseas program needs to be done "in order to maximise ensure⁹⁵ administrative efficiencies and congruence of practice", thus emphasizing the underlying issue of finances as well as the alignment of two colleges' practices in conformity with University priorities.

A relevant metaevaluation finding shared in the summaries in CS2, CS4 and CS5 is that the anonymous writer compares and contrasts the perspective of participants with that of

the Review Panel and offers an additional layer of interpretation of findings. For example, in CS2 the writer states that

It was clear that within the School, Faculty and wider University, there was recognition that [...]. The panel agreed that this was important but also felt that [...]

As another illustration, in the “Financial Management and Performance” section, the writer qualifies the financial position of the School as “*obviously strong*” [emphasis mine] which “impressed” the RP.

In CS4, the same pattern is found. For example, in the description of the finding from “Strategic Positioning of Language Study within the University” revealing that the program is split into two Schools, the writer contrasts the reviewers’ “concern” over this separation with the positive perception by members of the School, stating that “Teaching staff, however, indicated that there is a close liaison [...]”. In fact, this contrast is further emphasized in the following paragraph where while a new topic is introduced (e.g. the workload model in one of the Schools), how the language staff “saw many advantages in being located within” their particular School is restated. Therefore, this metaevaluation finding evidences the interpretation done by the summarizer.

Interestingly, in CS4 not only is the anonymous writer allowed to have the authority to take sides and emphasize certain findings, but also to disapprove of others and even, as mentioned, to remove information. For instance, after discussing in detail how the reviewers “considered” “dangerous” and “anomalous” the lack of “explicit acknowledgment” and the “omission” of the importance of languages from the “University’s Strategic Plan”, and the “cut off of language programs themselves”, the writer not only disapproves⁹⁶ of this because the RP had “ventured to highlight” such comments “At the risk of exceeding its remit”, but as interviewee (C) confirms, this section is left with no

recommendation since it was “excised from the final report”⁹⁷. But are there any consequences for this kind of action which compromises the authority of the reviewers’ voice? On the contrary, there is evidence in CS4’s Implementation Plan that before a plan is made, the recommendations need to be approved by the Vice Chancellor’s Committee. Therefore, this metaevaluation finding not only sheds light on the source of the decision to remove the recommendation, but also reveals that the main audience for the report are the commissioners who intervene in the review not only by procuring ToR but by deciding what is or what is not to be included in a “final report”⁹⁸.

Therefore, it is unclear how much of what is written reflects faithfully the reviewers’ findings and how much it is the result of the anonymous writers’ creation or as approved by the commissioners, which is not possible to test due to the lack of access to the reviewers’ actual report. Therefore, not only is the authority of the reviewers’ voice compromised but also the credibility of this summary.

Similarly, in CS5, while the anonymous writer introduces findings quite consistently with expressions such as “The panel noted that [...]”, “The panel was concerned at [...]”, “The panel believes that [...]”, thus revealing an attempt to establish a close connection between what is summarized and the original source, the writer also contrasts the view of the RP against the view of participants. For example, the positive view of a program requirement by the members of the School that “all Honours theses must be completed in the subject language rather than English” is contrasted with the different “opinion” of the RP⁹⁹. Moreover, sometimes it is uncertain whose judgment of merit and worth it is, as for instance when the summarizer states that “The Bachelor of Languages is *logically* the flagship program of the School. It was agreed that [...]” [emphasis mine].

3.7.2. Conclusions and Recommendations

3.7.2.1. CS2, CS4, and CS5

3.7.2.1.1. Conclusions

Neither CS4 nor CS5 offer a conclusion whereas in CS2, a “Summary/Conclusion” is presented at the end of the chapter which includes the anonymous writer’s concluding statements interpreting the RP’s positive outlook of the program. For instance, he states how “very impressed” the Review panel was with the School’s “current performance” since “there were many more positives” (e.g. diversity of programs; internal, national, and international reputation¹⁰⁰) than “negatives or challenges”. Moreover, the financial theme emerges once more in this section when pointing out further positives:

The School has been very well managed, and is in an excellent financial position. It has a positive and forward looking collegial culture and generally understands the need to carefully manage its programs in order to remain financially strong.

To end the conclusion, a diversity of the aspects “more worthy of consideration” are mentioned but which also include the need to conduct “reviews of those programs identified as potentially unsustainable in their current form”¹⁰¹ together with the “School’s strategic planning” which is also directly related to ensuring enrolment numbers do not decline.

3.7.2.1.2. Recommendations in CS2, CS4 and CS5

Since, as pointed out above, it is not possible to learn whether recommendations are based on the reviewer’s actual findings, metaevaluation findings are discussed taking into

account the extent to which recommendations are based on the summary of findings provided in CS2, CS4 and CS5.

In these three reports, recommendations are included at the end of every section in the Summary chapter which means these can be read as a conclusion.

In CS2, every section in the summary ends with commendations and recommendations which amount to thirteen (13) and seventeen (17) respectively, which denotes that while the RP's overall judgment of the School is considerably positive, several aspects require attention.

A description of positive findings is followed by "concerns" or "challenges" and suggestions for changes, and the commendations and recommendations ensuing logically from the findings is what is found in this summary. For example, in the Curriculum section, as described, the strengths highlighted, specifically, diversity of provision and approach, balancing diversity with sustainability and administrative support lead accordingly to six commendations praising (a) range of disciplines and "proactive approach" to sustaining diversity; (b) understanding of responsibility for viability of curriculum while acknowledging diversity of views as to how this may be approached; (c) accelerated mode of languages program reflecting the changing context of languages; (d) responsiveness of the administrative team, particularly of the School Manager, and focus on "enabling curriculum reform" via e-learning initiatives; (e) a School's committee in charge of promoting professional development through "educational innovation"; and f) discontinuation of Masters' coursework programs while maintaining units which contribute to other programs within the Faculty.

While the emphasis on commendations in general is diverse, and reflects concerns for the educational quality of the program, nevertheless a repeated emphasis on viability,

sustainability and marketing stands out. This can be discerned in the selected commendations and recommendations from the ES, in the brief description that precedes their listing in the ES and in the “Summary/Conclusion” section at the end of the chapter where the main results have been summarized.

As for challenges, concerns are associated with sustainability, as also described above and five recommendations are proposed: (a) to review a particular studies area to ensure ‘consistent’ representation of “particular countries” and “cross-listing” its units with other “relevant departments”¹⁰²; (b) improve balance of content and coherence of another studies area, which “would” “ideally” require an increase of one member of staff; (c) to review another studies area to make it more viable; (d) to consider the position of another studies area to “possibly” relocate it; and (e) develop a “strategic plan” to make the curricula of a number of areas “sustainable”.

In CS4, on the one hand, there is a clear correspondence between what is discussed in the summary and what appears in the recommendations, as in CS2. An example, as mentioned, is the limitations identified about “The Diploma of Languages” consisting of limited access¹⁰³, and lack of coordination reflected in the two recommendations which suggest strengthening cohesion, and promotion of the program, and removing “an inhibiting factor” that prevents students from enrolling. In “Promotion and Marketing of Language Study”, one of the three (3) recommendations, as a further example, proposes “An information day” based on “the low profile” of the University in marketing activities to promote language study.

On the other hand, however, sometimes issues raised in the discussion do not make it to a recommendation, as illustrated by the excised recommendation under “The strategic positioning of language study within the University” mentioned earlier. But additional

examples may be found. For instance, several issues brought up under “The organization and teaching of language courses” section (increase of staff’s contact hours not recognized in the work-model; limited number of students finishing with “a degree of language proficiency”), are not only not raised in a recommendation, but are diffused by emphasizing the positive side of language teaching. Likewise, while the discussion in “Language Proficiency and Benchmarking” comments on the difficulty of students in attaining proficiency which mainly arises from university constraints, the recommendation focuses solely on actions members of the program may take, such as benchmarking retention and attrition and supporting “high achieving students”; nothing is suggested to tackle constraints. As a further illustration, even though in “Staffing and Management” the situation is presented as serious across the languages due to increasing casualization and decreasing staffing, the recommendation chooses to focus only on one language area in need of staffing¹⁰⁴, likely in compliance with the ToR’s restriction to refrain from prioritizing recommendations requiring “extra resources”.

In CS5, before listing recommendations, a warning call is made about the focus of recommendations not being solely on “staffing issues”. Various possible implications may be inferred from this remark. On the one hand, it implies that the main audience of this report are those in control of resources, that is, the commissioners, and on the other hand, that ‘staffing’ may have become a regular theme in recommendations which, in turn, denotes a trend in resource deficiency in the area being evaluated. Lastly, it may also be inferred that commissioners, not willing to provide resources, become uninterested in reading the recommendations in the expectation of encountering requests on increasing resources, thus, using this comment to encourage their full attention.

Furthermore, as noted above, recommendations seem to aim at making the School comply with priorities. For instance, under “National and international research performance” R16 “to develop a more explicit and sophisticated approach to the embedding of research in the undergraduate curriculum” follows from University research priorities. Moreover, some priorities seem to be more urgent than others. For example, in “Capacity for further development of taught masters programs”, “urgency” is associated with the “little progress” made in the development of a “Masters of XX”, while the only recommendation under this section (R13) chooses to suggest the advancement of work in “the Master of Applied Linguistics and Master of Translation” “as a matter of urgency” in order to find “strategies” for “increasing enrolments”.

3.7.3. CS1, CS3, CS6 and CS7

Out of these four (4) reports, the only one with a designated section for a summary of findings is CS6. Additionally, it also has independent sections for a conclusion and recommendations. CS1 offers a section for recommendations and CS7, likewise, offers a summary following the Results chapter but rather than a summary of findings, it presents a list of the “Proposed Actions” placed at the end of every section in the Results chapter as the title “Summary of Proposed Actions and Implementation” suggests. However, no conclusion is included. In CS3, the information closest to this chapter expected by the ERC is a section named “Conclusion”.

3.7.3.1. CS6

In CS6, summaries, besides their inclusion in the ES, also appear both at the end of the main three sections in the Results chapter where the evaluation questions are addressed, as

well as after the Results chapter, and therefore are attributed special relevance. The summaries at the end of each section are clearly separated from the findings by providing a title (“Summary”) and a heading which expresses the issues that the review question asked to tackle with a key finding (“1. Summary: the past—on attrition, enrolments, timetabling *and surplus*”), or the overall aim of the question (“2. Summary: Proposals for future improvement”; “3. Summary: Increasing profitability”)¹⁰⁵.

While an extensive discussion and interpretation of findings is provided in the Results chapter as the self-reviewers announce in the ES, in the summaries it is minimal. Nevertheless, some commentary accompanies these bulleted summaries going from a superficial one (“This section is summarized below”) to more elaborate observations such as “In this section we addressed the question [...]”, and when appropriate a conclusion is also presented (“We have comprehensively demonstrated that [...]). Also, when appropriate, summaries are clarified. For example, after the four strategies “to further enhance profitability” are listed, a final comment points out a specific one, considering it to be “the most reliably effective”, that is, the reintroduction of higher units, including the rationale to justify this selection (e.g. minimal costs and a sizable surplus).

The “Concluding Statements” after the Results chapter comprise two sections. Their headings “Beyond finances” and “False review triggers and assumptions” clearly reveal a disapproval of the limited focus of the review and the biased grounds on which the review is framed. Accordingly, the first section highlights the program’s compliance with University goals as well as national strategic goals “beyond finances” (“strategic goals and internationalisation of education policy” which is then expanded in more detail emphasizing the alignment with “The Australian Federal Government” policies; “the extensive

connections with the outside world; and its “vision for future development [...] as demonstrated in this document”).

The contents in the second section exposes each of the triggers from the Review Notice (RN) with a commentary denying the underlying assumptions based on the findings from the Results chapter. This explanation in turn is introduced by a statement highlighting first how these triggers are perceived as “a concern”, and second stating compliance with the RN (“we have responded to the review panel’s questions¹⁰⁶”).

Two recommendations appear under a section before the Appendices (“Recommendations”), which as mentioned, are found too at the end of the ES, and therefore also highlighting their importance. These recommendations respond to the second task requested in the ToR. First, to make a recommendation about the future of the program, the self-reviewers provide a description of the program, which replicates the one in the ES, thus reinforcing the idea of the program’s compliance with University priorities (“an income generator”; “a crucial part of the University’s attractiveness to both local and international students”; “integral to and in strong alignment with the University Strategic Plan and International Education Policy”; “the main driver for exchange” in Asia; the program’s staff active role to responding “successfully to numerous changes over a short period of time”; to “extensive connections and partnerships in Australia and overseas” and to “marketing of the program and Faculty”) and it is presented as the basis for recommending the retention of the program. Additionally, with this recommendation, the authors not only challenge the false assumptions in the Review Notice but also stress that the arguments used for the pre-review decision to close down the program were unfounded.

Interestingly, the second task from the RN requesting the provision of “alternative options” for students to pursue their language studies in the state, is not addressed. Instead, the self-reviewers replace it with a second recommendation of a totally different nature, namely, the “improvement of information management systems”, thus highlighting a further criticism about the review process as seen in the concluding statements. The introductory note to this recommendation is based on findings commented on in the Results chapter in relation to deficient management practices and misinformed decisions affecting the program and the University alike, including the unjustified grounds for commissioning this review:

Accurate and timely information would have avoided the series of events that have led to the attempted closure of [language program], this unnecessary review, the internal and external reputation damage it has caused, and the staff time wasted at all levels.

While there is no evidence in the report, it may be possible to claim based on metaevaluation findings from CS6 so far discussed that this may be a conscious decision to protect the program since noting down future options for students could be interpreted as admitting that there is some truth in the assumptions made in the Review Notice and therefore that the program has been in fact badly managed by program staff.

Furthermore, while not under the recommendations section, the self-reviewers add recommendations regarding the second review question asking for strategies that can help strengthen the program on enrolments, retention, and costs (e.g. targeting the reduction of costs with the “most obvious option”, that is, “the reduction of face-to-face contact hours”).

3.7.3.2. CS3

The Results chapter in CS3 ends with a “Conclusion” that summarizes the “serious challenges” that the School and university-wide have had to face “over the last several years” ((a) establishing “a strong research identity”; (b) ensuring enrolments “remain strong”; (c) “carving out” a prominent place for language study “in the University’s educational priorities”; and (d) participating in Faculty initiatives involving “infrastructure, curriculum management, and academic workload”). This not only highlights the School’s success in aligning with these requirements but that compliance with priorities has made the School become a model within the University as the writer praises the School for leading “the way for other Schools”. Additionally, the final assessment the writer makes should be noted. While an eagerness to comply with University demands is stressed (“very important”), it is impossible not to detect at the same time an association of the expression ‘good Faculty citizens’ with ‘good behaviour’¹⁰⁷.

Nevertheless, to provide evidence of compliance, the writer restates in a bulleted list the twelve “major initiatives” described in the Results chapter as evidence and the three “challenges which the School has not “yet fully addressed”. An analysis of the initiatives reveals involvement in the *creation* of seven different programs; materials; or committees for the School (e.g. workload guidelines; an accelerated mode of the languages diploma; conference schemes; a technology innovation; an APD (Academic and Professional Development) committee; a new “studies” program collaborating with Asian, Chinese and Indian Studies; and an “in-country” study in Asia). The other five initiatives entail a variety of actions too (e.g. a review of course offerings; the digitalization of language lab library holdings; increasing the cohort of senior academics; the expansion of the Chinese

department; and maintaining “disciplinary identity while establishing collaborative links within the School.”).

Regarding the challenges that the School still faces, these are in connection with the “small size” of some departments; programs which “are far from realizing their potential” whereby “potential”, as discussed above, is mainly associated with enrolments; and lastly, the difficulty for staff to balance teaching and research and adapting to the research demands of the university. As far as recommendations are concerned, none are provided in this section though for some of these challenges some suggestions are made in the Results chapter (see p. 72).

3.7.3.3. CS1

In CS1, the main aim of the chapter following “Findings” is clearly suggested by its name (“Recommendations”). However, the recommendations in this section are not simply listed but preceded by an introduction providing a context that explains their purpose, a conclusion and a categorization of the recommendations. Hence, a four-folded purpose can be identified. Apart from reducing discontinuations, and increasing enrolments as suggested in the ES, the recommendations also aim at maximizing the students’ time while enrolled in the program and, additionally, the topic of cost-saving gets revealed as a purpose, when the reviewers express the view that the recommendations are also intended to minimize resources due to students who “enrol only to not continue”. To support these purposes, a rationale is also provided which serves, in turn, as a conclusion, that is, the acknowledgment of some unavoidable discontinuation, as also stated in the findings chapter, but the unacceptable percentage of current discontinuation (“far too high”). However, as also discussed above, the impossibility of finding out the level of acceptable discontinuation due

to lack of data to which this percentage could be benchmarked, and, as mentioned also in the findings chapter, the vulnerability of this degree to discontinuation given its ““ add-on” nature” are recognized as limitations of the study. Based on these claims, the reviewers suggest a specific “milestone” which involves cutting discontinuations by half “in the next few years”, thus highlighting the pressing need to tackle this issue. At the same time, the emphasis on cutting discontinuations, both for “keeping students” and “ideally” “increasing enrolments” implies that actions in the recommendations will focus on both purposes.

The introduction ends with an explanation about the division into urgent, administrative and academic recommendations, as pointed out in the ES, thus distributing the responsibility of the program among different stakeholder groups.

A closer look at these recommendations reveals that out of the twenty-four (24) recommendations, eleven (11) are justified in situ with actual data whereas the other thirteen (13) are not but which in the majority of cases are based on findings discussed in the Results chapter. In some cases, however, the reviewers reveal findings in the recommendations for the first time and use them to support a particular recommended action. For example, in encouraging the distribution of classes throughout the week, the reviewers justify it with data not mentioned in the Results chapter, namely, the lack of consensus in students’ responses about continuing to schedule classes at night.

Moreover, while the majority of the recommendations ensue from the findings based on the students’ own proposals, some recommendations are based on the reviewers’ theoretical and/or practical judgment. Many of these address educational concerns about program quality. For example, in the urgent recommendation about “placement”, the recommendation is justified by the reviewers’ strong criticism of the current placement criteria, judging it as “antediluvian at best and irresponsible at worst”, to then make two

different suggestions, one less demanding (“to test up” or “to test down”) and a more demanding one (“Computer-based diagnostic tests”). Likewise, the administrative recommendation on “Class sizes” is made on the basis of the reviewers’ judgment of considering classes of “20-30 students” “pedagogically disastrous” and indicating the “preferable” size of “15”. As another illustration, in the academic recommendation on the use of contemporary materials, the reviewers base their recommendation on the benefits they see in it, that is, because these “develop practical language abilities”.

In one particular case, that is, in recommendation 9 in the administrative section, a contradiction is found since while suggesting the promotion of overseas study, in the discussion of findings, the teachers’ “excellent job” at providing information about this is emphasized as well as that what students insist on having is more “on-campus” activities to practice the language. Therefore, a question arises about the reasons why overseas study is further encouraged by incorporating it into this recommendation. Is it perhaps perceived as a University priority that needs to be emphasized?

A particularly interesting example revealing efforts to squeeze into the recommendations suggestions to attract students is the “Team-teaching” recommendation under the administrative category. While this recommendation may contribute to the timetabling flexibility of teachers who can “mix and match” classes to suit their convenience, just as emphasized for students in the findings chapter, this recommendation is based on data not discussed in the Results chapter. In fact, increasing enrolments may very well be the priority associated with this strategy since in the first set of “urgent” recommendations, that is, the hiring of a program coordinator is advised whose role, among others is, as mentioned, to “run a major advertising campaign to increase [program name] enrolments”. The issue of finances emerges once more in this recommendation when the

reviewers assure that such hiring would not imply a high cost since it would “not require a full-time staff member”. Therefore, it could be claimed that underlying discontinuations and enrolments the effort to reduce costs also emerges. This is not to say that this is the only orientation in this report, where a concern for educational values is also evidenced.

3.7.3.4. CS7

Similarly, the summary of actions in CS7 is preceded by a very brief introduction revealing, on the one hand, a statement reassuring that the “recommended actions” are based on the review and the actions exposed in the Results chapter, which carry also a starting implementation date. On the other hand, the first topic mentioned in this brief statement states that “no program amendments are being proposed” with no further explanation. But what does “amendments” mean? Does this term imply that the actions proposed involve no significant program alterations, such as the reduction, restructuring, addition or discontinuation of program structures or programs? If so, this statement underlies the topic of resources and leads the reader to associate the main audience for this report with the commissioners.

A closer look at the Proposed Actions reveals four different categories according to the topic with which actions are connected. Moreover, each category includes a specific number of recommendations, namely, (a) Participation (R1-R4); (b) Employment prospects (R5-R10); (c) Retention and success (R11); and (d) Courses (R12). Most actions fall within Participation and Employment prospects.

But what does “participation” mean? The analysis of this category reveals an association with student numbers as well as student distribution. For example, the sub-headings “Enrolment numbers” and “First preference enrolments and TER cut off”

correspond to R1 which contains eleven (11) actions, thus making it the recommendation with the most actions, focusing on “the recruitment of domestic as well as international students”. The second part of the participation issue “Extending pathways and extending specialisations” involves R2-R4 with only one action each. In terms of distribution, a closer look at the eleven (11) actions under R1 allows us to see that distribution implies also the attraction of students, in particular, “higher TER students” but also of students of an aboriginal heritage as R4 seeks to study “the feasibility of providing” a major on aboriginal languages. The focus of the remaining ten (10) actions in R1, but also the remaining recommendations within this category are clearly geared towards attracting students or strategies for ‘keeping’ students which would have an effect on retention.

Interestingly, attracting students or keeping students is found not only in this first part of the summary but in many of the remaining recommendations and actions. For instance, in “Employment prospects” (R5-R10), R5 contains four actions and R6-R10 only one each, which aim at ensuring there is a close connection between what is offered in the program and career options. The actions under R5, in particular, emphasize the relationship with external stakeholders in promoting the program. In the same vein, the purpose of R6-R10 is to make the students in the program more ‘marketable’ and therefore actions entail promoting knowledge about the program through three strategies, namely, “graduate destinations”, incorporating a “career management information and skills into the program”, and “expanding the range of languages available”.

The remaining two recommendations, that is, R11 from “Retention and success” and R12 from “Courses”, furthermore, have to do with the investigation of the causes of ‘threats’ in connection with “low participation” in language courses at the University level and the impact of a state regulation which may affect enrolments. The nature of the action

in R11 is particularly similar to the ones under R1, that is, with monitoring retention to avoid a decline in enrolments since it proposes to “interrogate the retention and success rates” further.

Finally, while R12 focuses on the improvement of core educational content, what gets emphasized first is that the purpose is “to address [...] data” which as seen in the Results chapter, reflect “low satisfaction rates” in two courses. In fact, responses from student evaluations appear to be a key university priority since the first section of the Results chapter connects “the quality and richness of the student experience” with “University strategic emphases”. Therefore, given the widespread emphasis on attraction of students in this report, it may be claimed that “strategic” implies, in this case, a connection between student satisfaction and enrolments.

3.8. References and Appendices

3.8.1. Description

This is the last section in the ERC where references and Appendices are dealt with. For a list of references, a consistent and suitable style of format is suggested to ensure that references are free of errors and that they “cover all in-text citations”. As for appendices, the ERC highlights on the one hand the inclusion of “all appendices referenced in the text” and “in the order they are referenced”, and on the other, that “data and information in the appendices is clearly presented and explained”.

3.8.2. Findings

In terms of References, none of the reports have a list of references which in some cases such as CS2, CS3, and CS4 may be justified by the fact that no references are used at

all. Moreover, the absence of such a section in the rest of the reports may be due to the minimal number of research references used (and the absence of reference to professional documents as well, such as e.g. the Miron (2004) checklist).

Furthermore, all reports come with Appendices but in some cases, they have not been made available. For example, in CS5, the ToC shows the existence of Appendices, but none are attached at the end of the report and similarly the names of Appendices in CS3 are seen in the ToC but have not been made available for the purpose of this research. In CS6, out of the twelve (12) Appendices the last one has not been included which may be due to ethical reasons since it contained “48 pages of alumni letters”.

We may note the variation in the conception of what is and is not considered suitable information to go in an appendix and the decision to include or exclude certain type of information as an appendix. For example, information on submissions and meetings are made part of CS2 and CS4’s Appendices whereas CS5’s report section on ‘submissions’ is conceived as an internal section of the report outside the Appendices. Furthermore, while in CS4 there are two Appendices regarding submissions showing first, *invited* individuals, and other ‘groups’ of stakeholders to make a submission, and second, individuals *providing* a submission, in CS5, “Submissions Received”, as the title suggests, only tackles individuals *providing* a submission.

Moreover, in CS1, while the “introductory notes” from a quantitative report that was consulted composed by a University office are included in the one Appendix attached to the report, it is notable that no evidence is provided from another important quantitative study that was also consulted (“Some information was also taken from [...]’ 2002 report”) and which is pointed out by the authors as the grounds on which the commissioning of the current review is based,

It was envisioned as a qualitatively oriented, in-depth follow-up study to a report from the Faculty's Academic Programs office, prepared by [P34] in 2002.

of which only some general details may be found in the Results chapter:

The 2002 report was quantitative in nature and relied on data from the student administration database [...] and questionnaire responses.

In CS7, while an implementation plan is attached as an Appendix to the report, in CS4, the ToC clearly shows that the implementation plan which was made available for the purpose of this research is not integrated in the document as an Appendix.

As an additional example, while earlier metaevaluation findings proved that for CS2, CS4 and CS5 a self-review report was produced as a data source which in CS2 is made part of the Appendix section (e.g. CS3), in CS4 and CS5, it is not. Moreover, while CS6 as a self-review report commissioned within the context of a primary review clearly states its conceptualization as a 'submission', CS3 does not provide any explanation in a similar reference. Also, in CS6, the University did not grant access to the primary review report for the purpose of this research.

Therefore, several questions arise from these findings, for example, why are reports sometimes not made available? Why do some reports include a self-review in the Appendices while others do not? And why when Appendices, apart from a self-review report, do exist, they are sometimes not made available for research?

3.8.3. Appendices: Clearly Presented and Explained; Cited in Text; Attached in the Same Order as Cited

In the reports which include Appendices (e.g. CS1, CS2, CS4, CS6 and CS7), these are either partially or not explained at all with the exception of CS6. For example, CS1's

Appendix is quoted in the discussion of findings and the information it contains is clearly presented but its contents are not explored.

In CS2, out of the three (3) Appendices, only Appendix 3 is referenced as such in the Introduction to the summary of findings. Appendix 1 reveals the schedule of meetings with the RP, and Appendix 2 and 3 are “submissions”: the School’s “Self-Review” (i.e.. CS3), and a “Non-Confidential Submission”. The information is presented clearly in all of them, though in CS3, as explained, incomplete information and inconsistencies are found. Moreover, none of these Appendices have been explained although most of the information found in CS2’s summary of findings seems to rely generally on what is found in CS3. For example, the “accelerated mode for the program” is praised in various sections of CS3 (in “Management of Curriculum and Teaching” and “Innovation in Teaching and Learning Activities” where the program is presented as an example of the School’s commitment to “constant improvement in teaching and learning” and how it is “leading the way in innovative practices”; and under “Conclusion” as an illustration of “new and efficient ways of studying language in a modern university”). Likewise, this praising is consolidated in various sections in CS2. For instance, apart from highlighting in “Curriculum”, as noted, wider student access, it “reflects an acknowledgement of the changing context in which languages are being taught”. Similarly, in “School Positioning and Strategy” the program is emphasized as an “excellent example” of “awareness of wider educational opportunities that exist” associated with creating “distinctiveness” and, once more, “wider access and more opportunities for students”. Finally, in “Governance and operations” this program is also praised together with the School’s recent reviews for they “reflect well on the general leadership of the programs”.

As for Appendix 3 in CS2, its title “non-confidential submission” is followed by a brief statement which repeats verbatim the information that is learned in the Introduction

section of CS2 (“Three submissions were provided to the panel. Two of these were confidential. The third, non-confidential submission, is attached following.”) with no further explanation.

The contents in this Appendix reveal three paragraphs in a letter format carrying the signature of its author at the bottom of the document. The text starts with a subtitle indicating that it is a submission made “to the School’s review” and the name of a discipline area. The stated motivation of this submission is that the study area “is currently under threat” and the author uses the submission as an opportunity to make the collectors of submissions¹⁰⁸, who then supposedly pass them on to the RP, aware of the reasons why resources, particularly staffing, should be made available. The reasons alluded to for such a request are on the one hand, the increasing importance of the geographical region associated with this department in the world today which is further emphasized by stating that “engagement” with this part of the world is a priority within the university’s international plans. On the other hand, a second argument used is the fact that the teaching of this area has a tradition at this university which “was allowed to run down”, an idea which is further highlighted when commenting on the inappropriateness of the initial “resource rationale” for cross-subsidization. Therefore, what triggers this submission is a desire to explain a concern about the vulnerable financial viability of the program, to explain the reasons why this has happened, and also the reasons why this program should be financially supported, thus perhaps implying a suspicion that it may not be in the future. Interestingly, in CS2’s summary of findings, the concerns about this program are not based on finances but rather on the lack of expertise in this region and unevenness of offerings due to “research commitments of staff”.

3.8.3.1. CS3

In CS3, an analysis of the in-text references to Appendices reveals that some of the Appendices in the ToC (“Fee income summary”; the “Summer and Winter School”, the “Comparative Load”, the “Research Grant chart”, and the “Support Income”) are not cited anywhere. The rest of the Appendices are in the order they are listed in the ToC though a considerable number of irregularities and inconsistencies are found in relation to the numbering and/or denominations.

3.8.3.2. CS4

In CS4, Appendices are quoted in the “Review Process” section of the report and the contents of Appendix B are mentioned but not connected to it. Nevertheless, the information in the Appendices themselves is clearly presented and since they only provide evidence about submissions and meetings with the RP (Appendix A, B and D), and with documentation provided to the Review Panel (Appendix C), these are not referenced in the Results chapter, which reveals the role of these Appendices as offering evidence of procedures following policy requirements.

In Appendix A, the writer identifies, with no further explanation, the individuals who were ‘invited’ to make a submission though only partially. Hence, only the names of eleven individuals, mainly with high responsible administrative roles are provided, specifically five Executive Deans of different Faculties; one Deputy Vice-Chancellor; two Pro Vice-Chancellors; one Dean; one Vice-President; and one Acting General Manager whereas the rest of stakeholders are identified as ‘groups’ (“Heads of schools”, “Associate Deans”; the “student representatives and students of the program”; “members of the University community”; and “external stakeholders”). Therefore, while a variety of groups seem to

have been invited, the only concrete number of invitations that can be gathered is from the identified individuals in management positions. Moreover, the reason why only the names of these individual stakeholders are included is unclear, since no justification is provided. However, it may be possible to argue that this may respond to a policy requirement which becomes particularly interesting when Appendix B, which allows us to compare submissions received in contrast with invitations that were sent, reveals nine (9) submissions received, which do not correspond at all with any of the invitations from the named stakeholders in Appendix A. According to Appendix B, submissions number 1, 2, 3, 4, and 9 labelled with an acronym reveal an academic affiliation to this particular Faculty. Submissions number 5 and 6 carry the names of two (2) language students from two different languages. However, while in 5 the level of language is indicated (“L8 III”), in 6, it is not (“L6 Language student”). Submission 8 contains two (2) names and roles of external stakeholders representing a state government agency in connection with school education in general and to school education in languages in particular. Lastly, submission number 7 is more difficult to define. While the name and title of the submitter are clearly stated (“Adjunct Professor”), the academic expertise is not mentioned. Instead what is highlighted is the affiliation of this submitter to a company. Therefore, while an internal member of the university, there is a clear intention of associating this submitter with an external industry where, according to an online search, he is an Honorary Research Fellow. At the same time, while also identified as a “former student”, it remains unclear first whether this comment means to associate him as a former *language* student or not and second the extent to which being or not a former [language] student has any relevance in terms of submissions.

In Appendix C, a table provides clearly all the names of the different sources of “Information provided to the Review Panel”. This information is divided into seven main

categories, which may ensue from a review policy since no explanation is provided. They comprise (a) Introduction with three sections (the interview schedule; the Terms of Reference; and the Biographical Details of Panel Members); (b) Strategic Information which, shows University Plans at three (3) different levels; (c) Public Information (e.g. brochures on undergraduate studies (from University and Faculty), and printed website material (from Faculty and Program)); (d) Program Rules; (e) Comparisons, which reveals a “Benchmarking summary” with seven (7) universities from Australia; (f) the “School Submission Review”¹⁰⁹, which raises a question since the evaluand, according to the report, is housed in two different Schools; and (g) Statistics, which incorporates four (4) sets of statistical data, namely, an overview of both “Annual University Program Performance Reports” and “University’s submissions to XX” student data collection”; Students’ course experience questionnaires; graduate destinations surveys; and a variety of statistical data from one specific University campus including “enrolment numbers, load, gender, etc.”

Finally, Appendix D is a one-page interview schedule which carries the same title as the ToR¹¹⁰ and displays a detailed though unexplained table with information revealing times, places, and length of a total of thirteen interview sessions (ranging from 30 to 60 minutes) between the Review Panel and a variety of stakeholders, as the ToR instructed, during a two-day period. The titles, names and memberships of participants are clearly identified except in one case where only the name of the stakeholder is provided and which according to an online search is very well academically considered, in other words, an individual with an important research background, skilled in creating collaborative projects, and literate in Asia. In brief, as the table 8 below shows, a total of twenty-one participants can be identified, which reflects that a wide variety of stakeholders and also a variety of

languages were represented. Interviewee (C) emphasizes this diversity as a positive aspect of the review process:

C63. [...] we got to talk to a lot of a lot of different people including some people from neighbouring universities with whom (U3), (U3) had some agreements in relation to sharing language teaching.

The table too shows an imbalance in the number of students from different languages, and that no casual teachers are interviewed which is not justified anywhere in the report, as neither it is whether or not a Senior Lecturer and the students of Japanese did in fact participate since their attendance had not to be confirmed yet (“(tbc)”).

Participants		Languages
Administrators	4	
Academics with administrative roles	4	Chinese (1) Japanese (1) French (2)
Academics	4	Spanish (1) German (1) French (2)
External Stakeholders	3	HS Principals (2) State Agency (1)
Students	6	German (1) Japanese (5)
	21	

Table 8. Range of interview participants and languages represented in CS4.

Moreover, while not presented as an Appendix, CS4's "Implementation Plan" is described here since a similar plan is included as an Appendix in CS7, but most importantly because it extends the metaevaluation findings about CS4's review process.

CS4's Implementation Plan consists of a five-page document containing first a Memo letter indicating the names of the sender, the receivers, the subject concerned and a date, and second the plan of action for the thirteen recommendations ensuing from the review. The date indicates that the plan was finalized seven months after the review report date and the fact that it is "an updated implementation plan" means work during those months had produced at least one more version of this plan. The letter is sent by the Executive Dean of the Faculty to (a) an associate professor of a particular language area; (b) the manager of Strategy and Administrative services in that Faculty; and (c) to the two Heads of School within that Faculty in which the languages are located. Therefore, this is an internal memo from upper management to lower management, except for a senior professor but who, as Appendix D shows, acted as "Acting Executive Dean" during the review process.

As for the subject in the memo, it reads "Implementation Plan" followed by the exact title found in the ToR ("The Review of the Diploma in XX"). What follows shows a 4-page plan headed by the same title name and a brief introduction which reveal further the close connection of the commissioners/organizers of the review with the planning of actions ensuing from the review. In fact, the introductory paragraph preceding the plan evidences the different management roles in control of both the review and this plan, and within it, who is responsible for the preparation of the plan, in this case, the two Heads of School, and who for its implementation ("Quality and Reviews" to oversee the implementation process to ensure "appropriate progress" as well as that it "occurs on behalf of" the university)¹¹¹.

As far as the implementation plan itself, the information is presented in a clear way showing a five-column table, each one headed by the following categories: (a) Recommendation; (b) Action; (c) Designated Responsibility; (d) Date for Completion; and (e) Progress. All of the cells are filled with appropriate information except for the fifth column which means that this memo is the first step towards implementation. Moreover, the recommendations on this plan correspond exactly to the ones listed in the report. In terms of dates, most actions are given about a year-long to complete except for two (2) which are considered “ongoing”, although one of them has a date, too.

An analysis of the “Designated Responsibility” (“key accountabilities” in the memo letter) shows that the academics designated as responsible for actions and closer to the educational aspect of the program are Heads of language disciplines, program coordinators and the program convener. Interestingly, only in one instance does the program convener appear as the only one responsible, while the rest of the others appear “in collaboration with” someone in management. One apparent exception may be the Heads of language disciplines in one (1) Action but these are to be coordinated by a particular person whose name has been specified.

3.8.3.3. CS6

The information in the CS6’s Appendices is clearly presented but not all of them are explained. Furthermore, Appendices are not presented at the end in the order they are referenced inside the report.

The function of the twelve Appendices attached at the end of CS6 is to support findings in addressing the questions in the RN, as noted in the Introduction section of the report, and also to provide evidence in support of the program. Examples of this

complementary data are evidence of success from past reviews; information in relation to the internal recruitment strategies proposed by program staff (i.e. CS6's self-reviewers) under "Increasing student numbers issue"; revision of generic skills learning outcome, etc. Additionally, external and internal stakeholders' letters supporting the program are also attached as Appendices (e.g. from a High School's principal and from alumni).

3.8.3.4. CS7

In CS7, three very clearly presented Appendices are provided which are all cited and are positioned in the order they are referenced. Appendix 1 stands out for being referenced six (6) times in comparison with Appendices 2 and 3 which are referenced only once, thus denoting its relevance. Appendix 1's significance lies in that it shows that the Proposed Actions that ensue from the analysis in the Results chapter and summarized in Summary of Actions chapter have been integrated into a concrete plan, thus demonstrating measures have been taken in order to increase student numbers. In fact, the stated "Objectives" of the plan in the introductory paragraph substantiate the claim that enrolments are a key focus in this plan since they aim at increasing enrolments and TER "for the Bachelor of Arts". Moreover, since the focus on marketing in this report to attract students is due to, according to (R)'s verbal data, the need to improve the vulnerable area of the program, that is, viability, it must be assumed that student recruitment by staff is perceived as the main strategy to sustain the program financially and which adds a financial dimension to the role of academic staff at this university. Furthermore, the fact that the plan is "developed in collaboration" with a Division with expertise in "Marketing and Business Development Manager International" emphasizes the importance of aligning this plan with university guidelines which in turn sheds light on the commercial orientation at this institution.

As a matter of fact, the remaining contents and particularly the language used in the marketing plan support further the commercial orientation. For example, the plan is structured into (a) Analysis of “Measurables”; “Situational Analysis”, and “Market Opportunities” which targets two (2) kinds of students, namely, “Domestic” and “International”); (b) a summary table from “A Year in Review – Domestic Marketing Activity”, where appropriate information is provided about “Date”; “Target Audience”; “Marketing Activity”; “Outcomes” and “Comments” except for the six (6) cells out of eight (8) in “Comments” which are empty and about which no explanation or justification is provided. However, the comments provided in the other two (2) cells show that the purpose of this column is to provide verbal evidence that the marketing activity has been realized as suggested by the title in each of the two cells (“Some of the comments from the students include: [...]”); and (c) an additional summary table of “Future Planning- International and Domestic Marketing Activities for [program name], 2011-2012” which similarly contains the same four (4) columns as the previous table but adds a new one “Objectives” and leaves out “Outcomes”, logically, since these activities are part of “future planning”. The activities planned reflect the “Key Messages”¹¹² that summarize main topics from the analytical part of the plan and are organized into two timing blocks, one for 2011 and the second one for 2012 with an “Evaluation and Review of” the marketing program carried out thus far, indicating the ongoing character of this plan.

Appendices 2 & 3 consist of evidence complementing findings from the Results chapter, that is, as confirmation that findings reveal compliance with review policy to demonstrate alignment and compliance with University priorities since what is first presented is “Program Support”, demonstrating “collaborative activities and strategies” (e.g. a table with numerical data about students’ “reasons for contact” as evidence to prove

the level of collaboration between a “centre” and the program). In the same vein, a document prepared by an “academic librarian” is offered in Appendix 3 accounting for the Library services (“Library holdings”, “Books”, “Journals” etc.). Additionally, the “Expenditure for Languages and Linguistics holdings over the last 4 years” is also part of the information presented which is supported with a table with “expenditures” by the program which, while not labeled, is adequately discussed and explained.

Chapter 4: Discussion

Show us the profit, don't worry how you got there is a catch-cry in the financial pages. We see the accountability movement in the public sector as having been hijacked by economists, politicians and others who subscribe to the bottom line mentality, with the transfer of a methodological technology that fails to grasp the complexities of social interventions. (Owen 2004, p. 4)

Chapter 4: Discussion

4.1. Introduction

In this chapter and in the next, key items arising from the metaevaluation process of the CSs will be discussed and related to the Research Questions:

1. Is metaevaluation an effective method to find out the character of the review processes of language departments across Australia?
2. What is the character of the evaluations – is their focus mostly managerial or educative? What are the implications of this for the theory and practice of evaluation and metaevaluation?
3. How can the key findings of this metaevaluation help guide future LPEs align with good evaluation practices within a strongly ideological policy context?

4.2. Strategic Policy Use of Quality Assurance Evaluation for Accountability in Australian HE Preventing Improvement

Overall, the findings of this study confirm the value of metaevaluation in general as a tool for establishing the character of evaluative practice in language programs in Higher Education in Australia. The study demonstrates conclusively the primarily managerial character of the review processes of language departments across Australia. This finding echoes recent discussions of evaluative practice in Australian HE more generally. The reviews conducted at present on behalf of quality assurance within an individual Australian HE institution reflect their regulatory nature (Meyers 2012), also commented on by Massaro

(2013) and Coates (2013). It seems that whether at large or at small-scale, reviews are used as instruments of compliance. Meyers (2012) points at key characterizations of these reviews and some serious validity issues as the main cause that undermines their potential use for improvement. Firstly, there is the issue of using reviews to align objectives of a particular review with universities' broad goals such as "enhancing productivity, responding to student and employer demand, and widening access" (p. 5)¹¹³. Moreover, when the review is done at a small scale, such as a course, the review focuses on economic efficiency. To illustrate these tendencies, the report *Mapping the Humanities, Arts and Social Sciences in Australia* (Turner and Brass 2014) prepared by the Australian Academy of the Humanities, created as a not-for-profit organization to "advance knowledge of and the pursuit of excellence in the humanities in Australia", strongly suggests that program reviews are currently being used in Australian universities as a small-scale instrument to attain restructuring that in the past was achieved via "large scale institution-wide changes" both of which are associated with 'rationalize', which is defined as "make (a company, process, or industry) more efficient, especially by dispensing with superfluous personnel or equipment" ([https:// en.oxforddictionaries.com](https://en.oxforddictionaries.com)). The Academy report cites a report by the Australasian Deans of Arts, Social Sciences and Humanities (DASSH) entitled *Benchmarking the Australian Bachelor of Arts* (Gannaway and Sheppard 2012), which found that in the period prior to 2008 there 'was a fair degree of change and faculty restructuring' that impacted on the management of BA programmes across the country. In contrast, the Academy Report notes (Turner and Brass 2014 p. 31),

In the period 2008–12, however, 'only eight of the 35 institutions with a BA programme report[ed] a restructure' that had an impact. It is quite another story in terms of the changes to the BA programme itself, however. DASSH reports that in the period 2008 to 2012 there was 'substantial change' to 30 of 39 institutions under examination. Again, in the period prior to 2008, programme changes were essentially

the outcomes of 'institutional restructuring or large-scale institution-wide changes' whereas in the recent period to 2012 most changes are attributable to programme reviews. The DASSH report also foresees further likely rationalisation of majors 'particularly in the form of rationalising the number offered and the units offered within the majors' and anecdotal evidence over the last few years would certainly support this.

Meyers (2012) points out some important issues which may impact negatively on staff morale and can contribute to create a schism between those in management and those managed, preventing the necessary commitment to embrace reviews. While the organization or structure of a program may be in need of revision, Meyers (2012) refers critically to the constant endemic change and restructuring that takes place in universities today with no apparent "sensible advantage for the academics or by extension the students they teach" (p. 44- 45). Coady (2000) further adds that the "obsession with frequent appraisal or assessment of individuals and recurrent reviewing of departments has also added to the sense many academics have that, tenured or not, their jobs are on the line" (p. 22). The monitoring of academics' work in Australian universities seems to be on the rise:

[...] with the help of new technologies and a more aggressive approach to increasing academic efficient productivity, new bureaucratic mechanisms of control in support of academic performance at NPM-based universities have contributed to create a sense of permanent scrutiny and job insecurity at all levels and forms of academic performance, daily, weekly, annually and even in four-year cycles. (Martínez Marco and Pérez de León 2019, p. 98).

Meyers (2012) mentions a deterioration in the credibility of review processes by academic staff, due to repeated experiences of reviews not influencing practice, which indicates the standard use of reviews as a policy instrument and not for educational improvement. Therefore, reviews, next to other accountability requirements, are perceived as a bureaucratic burden which besides trivializing the academic dimension of a program,

department or school have “very little serious utility” due to their limited focus on administrative matters (Coady 2000, p. 17).

To manage a university is no easy task (Balderston 1989). The question is how departments of education and university managers alike, a few years from now, will justify their huge investment in the administrative machinery that comprises the present university bureaucracy, and their commitment to accountability and transparency requirements when data shows that this is at the expense of the deterioration of the quality of education of university graduates. This is a very complex setting in which Molony (2000) includes also academics, at least in part, as responsible agents: “The litany of our supine compliance in the face of tyranny is endless” (p. 82).

The ‘undeclared cool war’ (Coady 2000) and its effects sadly are increasingly less latent and more obvious to the eye¹¹⁴ and they, in an important way, undermine the benefits of the potential successes, as well as putting at risk the future of what could be a healthy and solid HE system in Australia. The response to reforms have been to “redirect spending from teaching and research to an ever increasing array of activities devoted to short-term revenue raising, public image, student recruitment and pointless compliance administration (Meyers, 2012 p. 33). Meyers also defends that “avoiding risk for the bureaucracy and minimizing cost of non-administrative activities is the only serious business [...] the constant employment insecurity invokes a general feeling of powerlessness to change anything for the better” (2012 p. 40, 144). In this context, it is unlikely to find the needed commitment by academics to evaluating programs for the purpose of raising the educational quality of programs.

However, the drive for accountability has not stopped as a result of the Bradley review, including discussions about the need of a national regulatory governmental body to

accredit universities based on standards. Various concerns have been raised with this approach since the Bradley report also reported that Australia lacked a data base on either student learning outcomes or system performance. Therefore, to address the demand of comparative data, a new framework needed to be created on the basis of these two issues. Currently, as Coates (2013) states there is in Australia “a significant amount of work [...] to develop policies and infrastructure for improving the assessment of student learning” (p. 32), the rationale for it being “an intrinsic interest in helping students learn and achieve.” (p. 33). However, Coates has identified issues that challenge the effectiveness of the whole institutional approach using performance measures and outcomes measures which the government agency known as the Tertiary Education and Quality Standards Agency (TEQSA) has been using since its creation in 2012. TEQSA was preceded by the no-longer operative Australian Universities Quality Agency (AUQA) and its main broad role is “to do something new: define and manage academic standards in Australian higher education.” (Marginson 2013, p.7). In fact, Coates (2013) argues that since Australia does not have at the moment the “institutional or intellectual architecture for an assessment framework to improve education” any attempt at measuring the performance of a particular institutional system may become the study of ‘provider standards’, that is, the measurement of the “measurable” rather than the measurement of what is important. In addition, quality may be defined in different ways, which poses a challenge when it comes to the elaboration of standards whose process should take into account contextual parameters for them not to run any risk of invalidity. Finally, if TEQSA’s role is to seek “information on quality and productivity to procure and justify increasing public spending” (Coates 2013, p. 33) an emphasis on accountability with TEQSA’s large-scale approach may put at risk the standards and quality that it seeks to protect (Massaro 2013).

For example, in explaining that “The current review grew out of” a meeting between the Dean and the Head of the program in which the Dean announced the closing down of the program which was then replaced by a “process of consultation and review”¹¹⁵, a political use of the review as a pretext to shut down the program is clearly suggested. In fact, CS6’s self-reviewers’ stated aim to address “some contentious and erroneous statements in the “Triggers” statement in the review notice” supports this suggestion.

The four triggers in the RN, while not exposed in the Introduction but in the “Concluding Statements” of CS6, allude to various specific problems about the program, namely, (1) “low enrolments overall in the major and the high attrition rate”; (2) “Financial cost to the faculty”; (3) “Lack of action towards implementation of the recommendations” and (4) “Community concern about earlier proposed closures”. The fourth trigger, which clearly refers to the cancellation of the program announced in the meeting explained above, is interestingly downplayed by the self-reviewers in the “Concluding Statements” stating that

Community concern about the proposed closure has focused on issues of educational and academic integrity, and procedural fairness; all of which are warranted.

whilst verbal data by interviewee (A) points out quite the opposite. Issues which affected not only the integrity of the teachers but also of the students are raised since the decision to close down the program was rushed and unexpected:

A2. [...] er in 2— in December 2011, and the dean proposed to close it in 2 months time. He said ‘oh, we will finish this year so we’re not gonna offer next year’ in December. It was already December 14 or something. Er no, it’s not even that, it was December— it was 20 or something. It was few days before Christmas.

Q3. And, and that meant that it wasn’t gonna be offered—

A4. Offered for a year. And er it was—

Q5. Just 2 months notice, something like that.

A6. And the students were already enrolling in. Students are choosing (U4), in some cases, because they can study something and (L6), and suddenly being told and there was no precursor whatsoever.

In fact, while (A) points out in the interview a stressful environment for language teaching (“Language was always under pressure”) which was due to “well, it’s a perception of ‘we are losing money’”, the announcement of the closure of the program was still perceived as “out of blue” since according to past reviews, program staff had not been made aware of any serious concerns:

A2. Okay er while I was at the (U4) which was from 2002 to 2013, we had quite a few reviews. Um ((sigh)) but the— we had a few reviews and each of them— actually review result wasn’t that bad. They said maybe we should make more effort to increase students in one way or another but our teaching curriculum was rated highly, student satisfaction was always high, so that was fine [...]

Furthermore, the Introduction of the report also reflects a surprising effect when the self-reviewers confirmed with the Dean that it was the financial aspect of viability which was the focus of the review since program staff had never been provided with financial outcomes reflecting issues nor were they ever made aware of an existing financial problem with the program:

A270. Okay, for instance, the important thing: Money, teaching course income issue. I have never been given the information and when I was told (L6) program has been losing a lot of money for many years, I was very surprised because no one said that. (L6) was breaking even, just we have to be careful (xxx).

In (A)’s view, this had been true in the past for some languages but unfounded at the current time, which is what contributes to perceive reviews as a threatening instrument (e.g. “when I joined the (U4) I already sensed all the language people are feeling that university is gonna come and get them, so review is used for that”).

Therefore, it may be claimed that the softening of the fourth trigger may have been used intentionally as a strategy to focus the attention on evidence about the other “false”

three triggers. In fact, this claim can be supported through the following quotes showing (A)'s conviction about the program's fine condition including finances and program staff meeting academic and research expectations:

A490. [...] So what I learnt was I ticked all of the box—

Q491. Yeah, that would be my last question.

A492. I ticked all the box. All pdf, I was never being questioned 'yeah, you're doing great, great, great' and the student number increased 3 times from 2004 to 2011. It was increasing.

A494. And our teaching evaluation was really high like usually 90%, 100% student satisfaction rate in every unit. That's how well we run it.

A496. Um so I'm very confident it wasn't the problem of the program. It was a very useful thing, financially not losing anything, even bringing in money and er (L6) language most widely taught language in (S/T1), and all the information, all (xxx) information I told them in manifold, and a lot of graduates write to the council and then saying how they thought— how useful that was to the [...]

In fact, the causes and the effect of the community concern raised in the fourth trigger are further explained by (A). It seems that it was due to the "campaigning", the "huge resistance, "the "media coverage" and the "noise" carried out by program staff and other program supporters in order to save the program after the decision to shut it down was communicated:

A214. Um actually the review happened 2012. 2011 was a notification of closure.

Q215. Closure of the whole languages program.

A216. Er 2011 is (L6) program, and then we made a lot of noise: Went on the media um yeah, brought in the media um also students did the online petition and they got a lot of names like professors from all over Australia, a lot of kind of—

Q217. Support.

A218. Support from influential people, as well.

A220. Influential people as well. And it was ta- (PP1) took that to the (S/T1) assembly so it's on the (xxx), it's a table. So we made a lot of noise, so they couldn't just unilaterally shut down as they wanted.

which caused "embarrassment" as suggested by (A), but also by interviewee (K), who claiming to have been involved in supporting program staff in this process, explains that the

decision to close down the program caused a “public outcry” and as a consequence embarrassed the commissioner:

K68. Well, it was the dean in this particular case, and the reasons why this dean did it this way— he was forced to do it as a result of public outcry because it became— the decision, the decision was made to close the language without discussion, then, there was public outcry in the press and he was forced to establish a review er to satisfy public concern and then— [...] — it was saved but with certain face saving er criteria for the dean.

In sum, the contextual information in CS6’s report as well as in the verbal data and the information in the RN is crucial in leading interviewee (A) to understand the commissioning of this review as “well, actually it wasn’t review of language program” thus suggesting the underlying use of the review as an excuse to make public the problematic viability of the program which according to the ES was “proven” to be erroneous. Therefore, in the end the review is seen as a political instrument to implement the dean’s decision (“So decision first”). In the same vein, (K) suggests that the review is used as a pretext not only to deal with finances but specifically to push a false argument, that is, a *deficit* in finances:

K64. [...] And one of the issues, of course, was the financial aspect of it. The driver for the review was, you know— it was supposedly costing money, losing money. Turned out, of course, it wasn’t losing money.

Moreover, besides the triggers the 3 questions from the Review Notice are disqualified by (A) as true questions since they are perceived as tendentious (“it’s all attacking, it’s not question”). As the following quotes show, in fact, not only the questions but the whole ToR are interpreted as prejudiced:

A6. [...] and the (L6) language program review was pretty outrageous: The review document um what was it, what’s called? The term of reference was all heavily critical, it wasn’t neutral, it was saying something like um (L6) program was, well, we had 5 reviews in the past and in all of them (L6) program was told to make various changes and (L6) program has been refusing to make any of those changes and then not in— and so forth.

The three evaluation questions are contained by the 4 “aims” stated in the Introduction of the report and which, as mentioned, intend to tackle several points: (a) to show compliance with the RN; (b) to prove wrong the concerns stated in the RN triggers; and (c) provide further justification in order to keep the program by discussing “the wider context”, that is, Government policy, as a main contributor “where it affects matters of future viability” .

In conclusion, the identification of the two lines of contending information, one from the Review Notice by the commissioners, and a second one by program staff, clearly shows that CS6 is an integral part of a larger primary review where explicit and implicit purposes can be identified for different stakeholder groups. On the one hand, the implicit purpose is to use the opportunity of the review by the commissioners to save embarrassment and then implement a pre-determined decision. Therefore, the public pressure did not have the desired accountability effect since the power of managers to implement false efficiency-based decisions overrides the evidence in the report. In fact, the ToR shown in the report further denote the initial firm decision to close the program when commissioners ask to provide recommendations about not only the future of the program but also of “possible alternative options for students” who wish to study the language in the same state. This evidence, in turn, creates the environment for a second implicit purpose for CS6’s self-reviewers to use the review to defend the program not only from the pre-review decision to close the program¹¹⁶ but also from the focus of the review on flawed information as the following statement from the ES highlights:

We have proven that the major triggers for the review were based on misinformation, and the premises on which the review questions are based are not accurate.

In CS3 and CS7, while no explicit reference is made about a purpose nor are any evaluation questions suggested, inference can help to shed light on these two evaluation issues in these case studies.

In CS3, since it is triggered by the initiation of a primary review (CS2), as in CS6, where it is to be used as a data source, its first purpose is compliance with regulation. However, evidence from the Introduction section in the report reveals a further possible purpose. For example, the School and the Faculty's "Organizational charts" mentioned in connection with structural changes that the School has undergone may be linked with the Strategic Plans mentioned in CS2's ToR. In fact, evidence in the program description in this same section showing that the School's research, service, curriculum development, language teaching and technology are to be "aligned with both the University and Faculty strategic plans" clearly leads to associate broader Strategic plans with the School's own scheme to meet requirements at Faculty and University levels in these five specific areas.

Interestingly, this does not seem to be the only Strategic plan to comply with since in "Management and Financial Performance" an additional broader plan is presented under the name of "University Economic Model" (UEM) and defined as "a financial management tool" designed for two specific purposes: first, for generating money ("to encourage sustainable business management at all levels of the organization") and second, for saving money (incentivize "budgetary discipline"), thus providing the "framework" for making programs, departments, Schools, etc. accountable for how money is spent (to achieve "greater transparency" in "resource allocation" and "costs"). Therefore, according to these findings, it can be claimed that demonstration of compliance at two levels (i.e. internal regulation and alignment of 5 specific areas with University strategic priorities following the financial principle ensuing from the UEM involving both generating and saving money) is the

main purpose in this case study. This raises two issues: first, to what extent is the program aligned with Strategic plans in connection with research, service, curriculum development, language teaching and technology; and second, to what extent does the alignment of these areas abide by the UEM guidelines?

With regards to CS7, the first section of the report “Purpose of the program and strategic considerations” denotes a very similar review purpose to the one in CS3. In fact, “current and future contribution of the program to University, Division, and School strategic goals” is the very first topic addressed in the report even before the purpose of the program is described. The description in this section thus emphasizes how the program “contributes” and “supports” broader goals at four levels, namely, (1) “University’s strategic emphases”; (2) the “University’s current goals 2010-2012”; (3) “the Corporate Plan 2010-2012”; (4) “the Division Plan 2010-2012”, and “the School’s priorities”, particularly in relation to (a) “teaching-research nexus” with the support of a specific research “Centre”; and (b) the education of students as “global citizens”. Therefore, since the establishing of the program’s directions on the stated predetermined strategic grounds is given a prevalence in the document, the main purpose of this review, and of this review report particularly, is to demonstrate the program’s results and efforts in the alignment of the program, particularly of the areas presented in the report (program’s aims and objectives, collaborations, demand and market position, program viability, retention and success, etc.) with such strategic priorities. The main evaluation question ensuing thus from this purpose is ,does this report show results and actions that lead to the alignment of the program with University priorities?¹¹⁷

In sum, the Strategic Plans in CS7’s report and equally the strategic “context” in CS2, CS3, CS4, and CS5 are a key element in these reports that may represent the criteria against

which the areas under review of the Schools or programs are examined for either demanding (CS2, CS4, CS5) or demonstrating (CS3, CS6, and CS7) compliance with University priorities. However, as noted, it is only in CS3 and CS7 where these are described only to a certain degree. Furthermore, while explicit evidence of an Economic Model guiding this alignment is only found in CS3, the earlier findings on the areas to be studied showing a general focus on student enrolments signals the relevance of finances in these reviews. Therefore, while not made explicit, a similar UEM to the one in CS2/CS3's university may be operated in the rest of the universities, too. Regarding CS1, while beyond student numbers, it also seeks to find out the reasons why students discontinue, its focus on enrolments reveals a concern with viability which in CS6, according to the evidence in the Review Notice, reflects an extrinsic interest in budgetary measures to keep the program financially viable.

Finally, in CS1, though the report does not contain Terms of Reference nor any evidence suggesting cycles of reviews, the "Brief and Scope" section states the existence of a previous review from which the current one is "a follow-up study". This thus suggests that the results from the first review may be the primary trigger for the current review.

Since no explanation is found in CS1 as to why the need to conduct a second evaluation using qualitative data in contrast with the previous quantitative oriented review, a potential reason could be a negative reception of findings from the quantitative report which may have prevented any commitment to actions (known in the evaluation field as evaluation utilization). While only further research could confirm this and also shed light on what may have caused such an opposition, the commissioning of this consecutive review suggests that the commissioners may have perceived the quantitative orientation of the first review as a potential trigger of the negative reaction, thus commissioning one based on

qualitative data. Moreover, two further evaluation issues were reconsidered and changed from the first to the second review as information in the “Brief and Scope” section indicates. The first one concerned the reviewers. Thus, while the identity of the “preparer” of the first report reveals an individual not affiliated to languages¹¹⁸, the membership of the reviewers in the present review reveals an area of expertise pertinent to the languages discipline (e.g. Applied Linguistics). And the second one concerned the source of the commission which changed from “the Faculty’s Academic Programs Office” in the first review to a closer-to-home commissioner, that is, the Head of School. Furthermore, the dates from both reports showing that there was a three-year lapse between the two reviews may too be an indication that perhaps no changes were implemented after the quantitative report and that the issue of discontinuation persisted. Interestingly, available anecdotal data suggests that the new methodological approach, the change of reviewers and of the commissioning source of the review were not issues affecting evaluation utilization since the findings from CS1 were not well taken either (“they hated it”)¹¹⁹.

This evidence, therefore, leads to the claim that the commissioning of the current review with the corresponding modifications may in fact have been triggered by the need to enhance evaluation utilization, that is, as an opportunity to gather evidence of a qualitative nature in an attempt to increase staff’s commitment to address the issue of discontinuations.

4.2.1. Academics vs. Managers

One of the findings of this study is the key involvement of managers vs. academic specialists in the evaluative process, which justifies the possibility that the implementation of a specific institutional agenda of action dominates the evaluative process. This finding should

be also considered in the context of a culture of increasingly lack of academic expert involvement in evaluative processes, being replaced by managers with a specific agenda (in many cases potentially related to institutional cuts). A greater question would be the separation of the remits of academic teachers and academic managers in order to ensure that teaching quality assurance processes are discussed in the context of managerial agendas with at least an equal weight, otherwise, the managerial agenda usually prevails in the evaluative process. The grouping of academics and managers also may confirm a lack of trust in academics' capability to work on their own or simply a lack of trust that academics would fulfill such tasks.

Placing the strategic information in second position, which in the ToR also have a prominent place, can be indicative that aligning the program with broad university goals constitutes the main grounds on which the purpose of the review is established. This inference is in fact confirmed by some interviewees, who get to associate university priorities with building a reputation within the country and staying competitive.

4.2.2. Managerial agenda implementation dominating the evaluation process

Findings in the case studies show that when a justification is provided for a particular recommendation, sometimes it is far from satisfactory. For example, in one of the Cs, after stating a positive finding about the School, it is mentioned that "The Panel noted the broad range of programs and courses", and then the writer states that the Panel, "too", "agreed that some reconfiguration is necessary", justified by stating that "the Bachelor of Languages" "is logically the flagship program" and therefore it needs to be "re-profiled as follows". Therefore, more than a justification, this sentence in reality underlies the need for

the program to implement several desired changes. In fact, this is not only confirmed by (A) later during the interview but it is the source of a complaint indicating unprofessional practices:

A8. No, it wasn't and there wasn't 5 reviews even. They, they don't even keeping track of number of reviews. [That's ridiculous.

However, the fact that the commissioners count this review among the ones in which the language area under review was involved may also indicate that program data may have been reviewed without the involvement and knowledge of program staff.

Behind the cyclical quality assurance trigger indicated in CS5's report there lies the political use of the review including the self-interested selection of reviewers to align the evaluand with specific University priorities which are associated with strategic goals such as research outputs and quality of language teaching while applying budgetary discipline which compromises the quality of language teaching, and making the program more easily accessible to make program competitive in order to attract students, thus increasing enrolments in line with NPM principles. This evidence reveals the commissioners' opportunistic use of the review and of the Review Panel to implant a particular agenda which evidently damages the integrity of academics, of the discipline, and contributes to an increasing antagonistic relationship between managers and staff. Furthermore, it also shows clearly that no matter the level of expertise in languages and awareness of the reviewers about the program's overt background disadvantages, choosing to participate in a review with organizational issues and an underlying agenda of imposition which is perceived as partial and unfair makes the reviewers to be perceived as agents serving the commissioners' interests.

Methodological shortcomings are also an indication of the priorities in the evaluation. To gather a full picture on methodologies in case study 4, what exactly can be found in the appendices of CS4 needs to be tackled, since they all seem to focus on data sources and collection methods. Findings on this appear later on in the metaevaluation report when dealing with Appendices but overall findings suggest that the information included in Appendices is used to demonstrate the transparency of the “Review Process” possibly in alignment with procedures from a policy which clearly leaves methodological rigor in a secondary place. For instance, a question that arises is whether participants in meetings were ‘invited’ to the meetings as was done for submissions, and if so, how they were selected, how data was collected, analysed, etc. Furthermore, these findings also denote a top-down policy approach to the review.

In summary, the heavy managerial emphasis in the commissioning, design and implementation of most of the program reviews studied suggests a failure to engage with the educational potential of program evaluation, which is the chief focus of the literature on this topic. While some aspects of the reviews address educational concerns, these are often subsidiary to or in the interest of financial and strategic concerns imposed from above. In the most extreme case, this is done from a defensive posture in the face of administrative financial plans at the expense of educational quality. Balancing evaluation priorities to give greater weight to the educational goals of an evaluation urgently needs to be addressed. Interestingly the relative neglect of the consideration of the prevalence of budgetary concerns in the program evaluation literature suggests the need for a theoretical correction as well.

Chapter 5: Recommendations and conclusion

Chapter 5. Recommendations and conclusion

This concluding discussion will focus on the third research question guiding this study, namely:

How can the key findings of this metaevaluation help guide future LPEs align with good evaluation practices within a strongly ideological policy context?

The discussion will consider this question in the light of broader discussions on the utilization of evaluation findings in program evaluation summarized in Chapter 1.

5.1. Recommendations

5.1.1 Utilization-Focused LPE evaluation

As we have seen in Chapter 1, evaluation utilization is what distinguishes language program evaluation from research on language programs (Alderson and Beretta 1992): the use of findings is the ultimate goal of evaluation: “an evaluation that is not used is in some important sense a failure” (Alderson 1992, p. 298). There is concern in the evaluation literature that the instrumental use of evaluation, that is, for decision-making and action, are less commonly observed than simply finding out what is going on in programs (their conceptual function) and being seen to be done to ensure their compliance with routine certification and (re-)authorization processes (their symbolic function) (Nevo 2009).

The strong managerialist orientation of the language program evaluations studied in this metaevaluation is evident in the findings and recommendations, which appear to conform to Nevo’s instrumental function rather than to a merely conceptual or symbolic one. That is, it seems likely that utilization of the findings will be an outcome of the evaluations. This instrumental focus is associated with the original initiation for the evaluations, which was typically top down and managerialist, often in the context of

financial constraints and the need to justify is possible the existence of the program in financial terms. While we cannot be sure what changes were made in programs as a direct result of the evaluations studied in this thesis, as this was not part of the research, the clear evidence of the context of New Public Management ideology in universities in the data suggests that the recommendations of the reviews, insofar as they focused on managerial and financial issues, were likely to be implemented.

Many of the findings and recommendations of the reviews reflect these managerialist concerns, as well as in some cases educational ones. In CS1, recommendations are divided into urgent, administrative and academic categories, but all seen in the light of the larger purposes of the evaluation which have to do with enrolments and costs. The findings of CS2 focus on the extent to which departments within the School are viable and serve the University strategically, and also whether strategies have been applied to align them with University priorities, particularly to make them sustainable. The program which is the subject of the evaluation in CS3 is generally praised, often in terms of its compliance with university priorities, making it a model for other schools in the university to follow. What challenges the program faces involve viability, enrolments, and adapting to the research demands of the university. In CS4, there is a focus on marketing, student enrolments and finances. In CS5, the recommended changes such as greater flexibility, while potentially an educational matter, are clearly associated with enrolments, and other recommendations reflect issues of funding and compliance with University priorities. In CS7, attracting or keeping students and adherence to strategic priorities.

As far as the recommendations reflect an educational focus, the anxiety in the utilization literature on the uptake of recommendations is perhaps more realistic. In order to achieve uptake, Alderson (1992) argues for anticipating the question of utilization in the design of the evaluation, to prevent the evaluation report and its recommendations languishing on a shelf somewhere. The key feature in the design of an evaluation to ensure utilization of its findings is to involve the stakeholders in negotiating the design of the evaluation. This was clearly seldom the case in the evaluations studied in this thesis. On the contrary, the top-down, managerial character of the evaluations put the teaching staff on the defensive. This is most clearly revealed in CS6, the self-evaluation which was used as input to a broader evaluation of the language program under review. This review was unnecessary, according to the staff: 'Accurate and timely information would have avoided the series of events that have led to the attempted closure of [the language program], this unnecessary review, the internal and external reputation damage it has caused, and the staff time wasted at all levels.' The self-review focuses on educational matters only in the light of the program's compliance with university priorities – the generation of income, attracting students, including international students, alignment with university policies, marketing, international alliances and so on.

The top-down focus of the evaluations means that the utilization-focused model of evaluation based on democratic principles developed by Patton (2008) and called by Norris (2008) the most influential evaluation approach of the 20th century is nowhere to be seen. Patton's 'process use' dimension of evaluation utilization, the learning that individuals gain through participation in the evaluation, is also hardly addressed in these reviews, although the frustration evident in CS6 is hopefully not typical.

Ironically, then, the concerns of evaluators that the conceptual (for understanding) and symbolic (for compliance) uses of evaluation are more prevalent than the instrumental (for decision-making and action) use of evaluation (Nevo 2009) appeared not to be true in the evaluations studied in this thesis. Here, practical, policy-based decision making, including the implementation of cuts to programs, was the focus. The studies were essentially managerial and motivated by specific and sometimes urgent policy and financial demands rather than representing routine symbolic compliance. The conceptual and educative function of the evaluations was backgrounded. It is recommended on the basis of the findings of this thesis that a better balance between educative and managerial aspects of language program reviews be achieved.

We have seen in the literature review in Chapter 1 that empirical studies in Language Program Evaluation showcasing good evaluation practices that aimed at utilization show how this might be done. As we have seen, Watanabe et al (2009) provide a survey of a number of evaluations in which there was local engagement in every step of the evaluation process including acting on the findings.¹²⁰ A summary of those evaluations is provided in Table 1 in Chapter 1. We refer the reader to the discussion in that chapter of ways of ensuring utilization of findings in a more educationally focused evaluation.

Overall, there is a gap between the findings of this thesis and the recommendations for utilization found in the literature. This reflects a broader gap: the design of the evaluations appeared to pay little attention to the literature on program evaluation in general, and language program evaluation in particular. It is recommended that in future evaluations of Australian university language programs, more time and resources be

devoted to the involvement of all stakeholders in the design of the study, and a greater balance between managerial and educational focus be achieved. The literature on program evaluation and language program evaluation has much to offer here.

However, there is a weakness in the program evaluation literature which is revealed by the case studies examined in this thesis. The harsh political and financial realities of program management, and the competing interests of stakeholders, are only weakly reflected in the literature, which tend to take a more idealistic stance on evaluations, focusing more or less exclusively on their educational character. It is recommended that future theorizing about how language program evaluation can be implemented find a better way of balancing educational and managerial imperatives, particularly in the current era of New Public Management in university education.

5.1.2. Limitations of the study: Criteria for metaevaluation

A vast amount was learned about conducting a metaevaluation in the course of the completion of this study. Reflecting on its conduct, there are a number of ways in which it succeeded, and a number of ways in which it might have been improved. Establishing criteria against which a metaevaluation will be judged ideally should be done in advance of the metaevaluation. Such criteria would involve: (1) Appointing qualified evaluators as part of the metaevaluation team; (2) identifying and fostering interaction with the stakeholders in the metaevaluation; (3) defining the questions to be addressed; (4) agreeing on criteria for judging the evaluation system or single evaluation; (5) ensuring there is a formal contract or a memo of understanding for the metaevaluation; (6) collecting and reviewing relevant information; (7) collecting new information as needed; (8) analyzing and synthesizing findings; (9) judging the adherence of the evaluation to appropriate standards

and/or criteria; (10) disseminating the findings of the metaevaluation via a variety of means; and (11) assisting stakeholders to interpret and apply findings. These criteria were partly met in this metaevaluation: the most obviously missing aspects are (1) and (5). (2) and (7) were done to the extent possible. These limitations were as a consequence of the metaevaluation constituting an academic thesis rather than being commissioned by stakeholders. (10) and (11) it is hoped will be achieved by the availability of this thesis. The complete set of criteria would be recommended for future language evaluations as key criteria for conducting the metaevaluation.

5.2. Conclusion

This metaevaluation has contributed to highlighting the misalignments of Australian reviews with quality evaluative principles, mainly due to a policy context affected by principles not aligning with evaluation quality assurance. Thus, questioning and improving context in its alignment to good evaluation practices should inform future Australian LPE metaevaluations. Furthermore, it could be one of the main focus of all LPE metaevaluators to contribute to the improvement of the educational quality language programs. Moreover, in thinking about language program evaluation, the need to successfully combine the educational and educative goals of the evaluation need to be an ongoing priority.

Works Cited

- Alderson, J. C. (1992). Guidelines for the evaluation of language education. In J. C. Alderson and A. Beretta (eds.), *Evaluating Second Language Education* (pp. 274-302). Cambridge: Cambridge University Press.
- Alderson, J.C. and Beretta, A. (eds.) (1992). *Evaluating Second Language Education*. Cambridge: Cambridge University Press.
- Alkin, M. C. (2011). *Evaluation essentials: From A to Z*. New York, NY: The Guilford Press.
- Alkin, M. C. (1975). Evaluation reflections. Evaluation: Who needs it? Who cares? *Studies in Educational Evaluation*, 1(3): 201-212.
- Alkin, M. C. (2003). Evaluation theory and practice: Insights and new directions. In C. A. Christie (ed.) *The Practice-Theory Relationship in Evaluation* (pp. 81-89). (*New Directions for Evaluation*, 97). San Francisco, CA: Jossey-Bass.
- Alkin, M. C. (2004). Rooting for evaluation: A Cliff Notes version of my work. In M. C. Alkin (ed.) *Evaluation Roots: Tracing Theorists' Views and Influences* (pp. 154-169). Thousand Oaks, CA: Sage.
- Alkin, M. C., C. A. Christie & A. T. Vo (2013). A wider roots perspective. In M. C. Alkin (ed.) *Evaluation Roots: Tracing Theorists' Views and Perspectives* (pp. 386-393).
- Australasian Evaluation Society. *Guidelines for Ethical Conduct of Evaluations (1997)*
<https://www.aes.asn.au/join-the-aes/membership-ethical-guidelines/7-aes-codes-of-behaviour-ethics.html>. April 20, 2020
- Balderston, F.E. (1989). Academic program review and the determination of university priorities. In M. Kogan (ed.) *Evaluating Higher Education* (pp. 122-133). London: Jessica Kingley Publishers.
- Barber B. (1988). *The Conquest of Politics*. Princeton: Princeton UP.
- Bernhardt, E. (2006). Student learning outcomes as professional development and public relations. *The Modern Language Journal* 90(4): 588- 590.
- Bexley, E. (2013). On the fragmentation or decline of academic work. In S. Marginson (ed.) *Tertiary Education Policy in Australia* (pp. 97-104). Melbourne: Centre for the Study of Higher Education.
- Birckbichler, D. W. (ed.) (2006). *Evaluating Foreign Language Programs*. Columbus, OH: The Ohio State University.

- Bleiklie, I., J. Enders, B. Lepori and C. Musselin. (2011). New public management, net-work governance and the university as a changing professional organization. In T. Christensen and P. Laegreid (eds.) *The Ashgate Research Companion to New Public Management* (pp. 161-176). New York: Routledge.
- Boston, J. and Boston J. (1996). *Public Management: The New Zealand model*. Auckland: Oxford University Press.
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal* 9(2): 27-40.
- Braun V. and Clark, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology* 3(2): 77-101.
- Brown, J. D. (2012) EIL curriculum development. In L. Alsagoff, S. McKay, G. W. Hu, & W. A. Renandya (Eds.). *Principles and Practices for Teaching English as an International Language* (pp. 147-67). London: Routledge.
- Bryman A. (2004). Triangulation. In M. Lewis-Beck, A. Bryman, and T. Futing Lieao, *Encyclopaedia of Social Science Research Methods*. Thousand Oaks, CA: Sage.
- Bustelo, M. (2003). Evaluation of gender mainstreaming: Ideas from a meta-evaluation study. *Evaluation* 9(4): 383-403.
- Bustelo, M. (2006). The potential role of standards and guidelines in the development of an evaluation culture in Spain. *Evaluation* 12(4): 437-453.
- Butler, L. (2004). What happens when funding is linked to publication counts? In H. Moed, W. Glanzel and U. Schmoch (eds.) *Handbook of Quantitative Science and Technology Research* (pp. 389-405). Dordrecht: Kluwer Academic.
- Byrnes H. (2008) From representation at the federal/national level to creating a foreign language education framework. *The Modern Language Journal*, 92, 614-20.
- Byrnes, H. (2006). Introduction to *Perspectives*. *The Modern Language Journal* 90(4): 574-576.
- Calvin, L. M. and Rider, N. A. (2004). Not your parents' language class: Curriculum revision to support university language requirements. *Foreign Language Annals* 37(1): 11-25.
- Carr, W., & Kemmis, S. (1986). *Becoming Critical: Education, Knowledge and Action Research*. London: Falmer.
- Caulley, D. N. (1983). Document analysis in program evaluation. *Evaluation and Program Planning* 6: 19-29.

- Cisneros-Cohernour, E. J. (2012). Case studies and validity. In C. Secolsky and D.B. Denison (eds.) *Handbook on Measurement, Assessment, and Evaluation in Higher Education* (pp. 510-517). New York: Routledge.
- Coady, T. (2000). Universities and the ideals of inquiry. In T. Coady (ed.) *Why Universities Matter: A Conversation about Values, Means and Directions* (pp. 3-25). Sydney: Allen and Unwin.
- Coaldrake, O. P., & Stedman, L. (1998). *On the Brink: Australia's Universities Confronting their Future*. St. Lucia, Qld: University of Queensland Press.
- Coates, H. (2013). Assessing higher education outcomes and performance. In S. Marginson (ed.) *Tertiary Education Policy in Australia*, 31-38. Melbourne, Victoria: Centre for the Study of Higher Education.
- Cook, T. D. & Gruder, C. L. (1978). Metaevaluation research. *Evaluation Review* 2(1): 5-51.
- Cooksy, L. J. & Caracelli, V. J. (2009). Metaevaluation practice. *Journal of Multidisciplinary Evaluation* 6(11): 1-15.
- Cooksy, L. J. & Caracelli, V. J. (2005). Quality, context, and use: Issues in achieving the goals of metaevaluation. *American Journal of Evaluation* 26(1): 31-42.
- Cousins J. B. and Ryan, K. E. (2009). Dilemmas for educational evaluation in a globalized society. In K.E. Ryan and J.B. Cousins (eds.) *The SAGE International Handbook of Educational Evaluation* (pp. 539-555). Thousand Oaks, CA: Sage.
- Cronbach, L. J. (1963). Course improvement through evaluation. *Teachers College Record* 64: 672-683.
- Cronbach, L. J. & Shapiro, K. (1982). *Designing Evaluations of Educational and Social Programs*. San Francisco, CA: Jossey-Bass.
- Cronbach, L. J., S.R. Ambron, S.M. Dornbusch, R. D. Hess, R.C. Hornik, D.C. Phillips, D.F. Walker and S.S. Weiner (1980). *Toward Reform of Program Evaluation*. San Francisco: Jossey Bass.
- Cuthbert, R. (2011). Failing the challenge of institutional evaluation: How and why managerialism flourishes. In M. Saunders, P. Trowler, and V. Bamber (eds.) *Reconceptualising Evaluation in Higher Education: The Practice Turn* (pp. 133-138). Maidenhead: McGraw- Hill.
- Dahler-Larsen, P. (2006). Evaluation after disenchantment? Five issues shaping the role of evaluation in society. In I. Shaw, J. C. Greene, & M. M. Mark (eds.) *The SAGE Handbook of Evaluation* (pp. 141-160). London: Sage.

- Dahler-Larsen, P. (2009). Learning-oriented educational evaluation in a contemporary society. In K.E. Ryan and B.J. Cousins (eds.) *The Sage International Handbook of Educational Evaluation* (pp. 307-323) Los Angeles: Sage.
- Darling-Hammond, L. (1990). Teacher evaluation in the organizational context. *Assessment of Teaching: Purposes, Practices, and Implications for the Profession*.
<http://digitalcommons.unl.edu/burosassessteaching/7> April 7, 2020.
- Darling-Hammond, L., Wise, A.E. and Pease, S.R. (1986). Teacher evaluation in the organizational context: A review of the literature. In E.R.House (ed.) *New Directions in Educational Evaluation* (pp. 203-253). London: The Falmer Press
- Dassier, J. P. & Powell, W. (2001). Formative foreign language program evaluation: Dare to find out how good you really are. Dimension 2001: The odyssey continues. In: *Selected Proceedings of the 2001 Conference of the Southern Conference on Language Teaching* (pp. 15–30). Birmingham, AL.
- Dassier, J.L, Kawamura, H., & Costner, K. (2006). Using focus groups effectively. In D. Birckbichler (Ed.) *Evaluating foreign language programs*. (pp. 63-79). Columbus, OH: The Ohio State University.
- Datta, L. (2006). The practice of evaluation: Challenges and new directions. In I. Shaw, J.C. Greene, and M.M. Mark (eds.) *The Sage Handbook of Evaluation: Policies, Programs and Practices* (pp. 419-438). London: Sage.
- Fitzpatrick, J. L., Sanders, J. R. & Worthen, B. R. (2011). *Program Evaluation: Alternative Approaches and Practical Guidelines*. (4th ed.) Boston, MA: Pearson.
- Flyvbjerg, B. (2011). Case study. In N. K. Denkin and Y. S. Lincoln (eds.) *The Sage Handbook of Qualitative Research* (pp. 301-316). Thousand Oaks, CA: Sage.
- Gannaway, D., & Sheppard, K. (2012) *Benchmarking the Australian Bachelor of Arts*. Australasian Council of Deans of Arts, Social Sciences and Humanities.
<https://apo.org.au/node/69010>
- Gorsuch, G. (2009). Investigating second language learner self-efficacy and future expectancy of second language use for high-stakes program evaluation. *Foreign Language Annals* 42(3): 505–540.
- Gowing, D.B. and Millman, J. (1978). Can meta-evaluation give a direction for research on evaluation? *Annual Meeting of the American Educational Research Association*.
<https://eric.ed.gov/?id=ED161938> . April 20, 2020.

- Grayson, T. E. (2012). Program evaluation in higher education. In C. Secolsky and D.B. Denison (eds.) *Handbook on Measurement, Assessment, and Evaluation in Higher Education* (pp. 459-472). New York: Routledge.
- Halligan, J. (2011). NPM in Anglo-Saxon countries. In T. Christensen and P. Laegreid (eds.) *The Ashgate Research Companion to New Public Management* (pp. 83-96). New York: Routledge.
- Hammersley, M. (2012). Transcription of speech. In S. Delamont (ed.) *Handbook of Qualitative Research in Education* (pp. 439-445). Cheltenham: Edward Elgar Publishing.
- Harman, G. (1989). The Dawkins Reconstruction of Australian Higher Education. *Higher Education Policy* 2(2): 25-30.
- Harman, G. (1998). Quality assurance mechanisms and their use as policy instruments: Major international approaches and the Australian experience since 1993. *European Journal of Education* 33(3): 331-348.
- Harman, G. (2006). Adjustment of Australian academics to the new commercial university environment. *Higher Education Policy* 19(2): 153-172.
- Havelock, R.G. (1971). The utilization of educational research and development. *British Journal of Educational Technology* 2 (2): 84-98.
- Henry, G. T and Julnes, G. (1998). Values and realist evaluation. In G.T. Henry, G. Julnes and M.M. Mark (eds.) *Realist Evaluation: An Emerging Theory in Support of Practice (New Directions for Evaluation, 78)* (pp. 53-71). San Francisco, CA: Jossey-Bass.
- Hill, K. & McNamara, T. F. (2003). Assessment research in second language curriculum initiatives. In J. P. Keeves and R. Watanabe (eds.), *International Handbook of Educational Research in the Asia-Pacific Region* (pp. 629-640). Dordrecht: Kluwer Academic.
- Jacobson, P. L. H. (1982). Using evaluation to improve language education. *The Modern Language Journal* 66(3): 284-291.
- Jarvis, G. A. and Adams, S. J. (1979). Evaluating a second language program. *Language in Education: Theory and Practice* 19. Arlington, VA: Center for Applied Linguistics.
- Johnson, D. M. (1992). *Approaches in Research in Second Language Learning*. New York and London: Longman.
- Karlsson, O. and Conner, R. F. (2006). The relationship between evaluation and politics. In I. Shaw, J.C. Greene, and M.M. Mark (eds.) *The Sage Handbook of Evaluation: Policies, Programs and Practices* (pp. 225-240). London: Sage.

- Kawamura, H. (2006). Program evaluation as ethnography. In D. Birckbichler (ed.), *Evaluating foreign language programs* (pp. 15-28). Columbus, OH: The Ohio State University.
- Kiely, R. (2009). Small answers to the big question: Learning from language programme evaluation. *Language Teaching Research* 13(1): 99-116.
- Kiely, R. and Rea-Dickins, P. (2005). *Program Evaluation in Language Education*. New York: Palgrave MacMillan.
- King C. and James, R. (2013). Creating a demand-driven system. In S. Marginson (ed.) *Tertiary Education Policy in Australia* (pp. 11-20). Melbourne, Victoria: Centre for the Study of Higher Education.
- Kogan, M. (ed.) (1989). *Evaluating Higher Education*. London: Jessica Kingsley Publishers.
- Krause, K. L. (2011). Whole-of-university strategies for evaluating the student experience. In M. Saunders, P. Trowler, and V. Bamber (eds.) *Reconceptualising Evaluation in Higher Education: The Practice Turn* (pp. 139-144). Maidenhead: McGraw- Hill.
- Kushner, S. (2000). *Personalizing Evaluation*. London: Sage.
- Kvale, S. (2009). *Interviews: Learning the Craft of Qualitative Research Interviewing*. Thousand Oaks, CA: Sage.
- Lo Bianco, J. (1994). *Unlocking Australia's Language Potential: Profiles of Languages In Australia*. Canberra, ACT: National Languages and Literacy Institute of Australia.
- Lo Bianco, J. and Slaughter, Y. (2009). *Second Languages and Australian Schooling*. Camberwell, Victoria: Australian Council for Educational Research.
- Lynch, B. K. (1996). *Language Program Evaluation. Theory and Practice*. Cambridge: Cambridge University Press.
- MacDonald, B. (1987). Evaluation and control of education. In R. Murphy and H. Torrance (eds.) *Issues and Methods in Evaluation* (pp. 36-48). London: Paul Chapman.
- MacDonald, B. (1977). A political classification of evaluation studies. In D. Hamilton, D. Jenkins, C. King, B. MacDonald and M. Parlett (eds.) *Beyond the Numbers Game* (pp. 224-227). London: MacMillan.
- Mackay, R. (1994). Undertaking ESL/EFL programme review for accountability and improvement. *ELT Journal* 48(2): 142-149.
- Mackay, R. (1988). Program evaluation and quality control. *TESL Canada Journal* 5(2): 33-42.
- Marginson, S. (2009). Tradition and change in universities: The case of Australia. Paper presented at Hijiya University organized by the Research Institute for Higher Education

(RIHE), Hiroshima University, Japan Centre for the Study of Higher Education (CSHE), and the University of Melbourne, Australia. Retrieved from <https://melbourne-cshe.unimelb.edu.au/about/fellows/simon-marginson>

- Marginson, S. (2010). Trends in the internationalisation of higher education: The new dynamism of global action. Paper prepared for the internationalization of Higher Education Conference in Madrid, Apr 19–20 2009. Retrieved from <http://www.oapee.es/dctm/eu2010/19y20deabril/conferenciasesimonmarginson.pdf?documentId=0901e72b800d7f99>
- Marginson, S. (2013). Introduction. In S. Marginson (ed.) *Tertiary Education Policy in Australia* (pp. 7-10). Melbourne, Victoria: Centre for the Study of Higher Education.
- Mark, M. M. (2008). Building a better evidence base for evaluation theory. Beyond general calls to a framework of types of research on evaluation. In Smith, N. & Brandon, P. (eds.) *Fundamental Issues in Evaluation* (pp. 111-134). New York: Guilford Press.
- Mark, M. M., Greene, J. C. and Shaw, I. F. (2006). The evaluation of policies, programs, and practices. In I. F. Shaw, J. C. Greene, and M. M. Mark (eds.), *The SAGE Handbook of Evaluation* (pp. 1-30). London: Sage.
- Mark, M. M., Henry, G. T. & Julnes, G. (1999). Toward an integrative framework for evaluation practice. *American Journal of Evaluation* 20(2): 177-198.
- Mark, M. M., Henry, G. T. & Julnes, G. (2000). *Evaluation: An Integrative Framework for Understanding, Guiding, and Improving Policies and Programs*. San Francisco: Jossey Bass.
- Martín, M. D. (2005). Permanent crisis, tenuous persistence: Foreign Languages in Australian universities. *Arts and Humanities in Higher Education* 4(1): 53-75.
- Martínez Marco, L. A. (2013). *Evaluating Program Evaluations: The character of language program evaluations in Australian language departmental reviews*. MA Minor Thesis, School of Languages and Linguistics, The University of Melbourne.
- Martínez Marco, L. A. and Pérez de León, V. (2019). Metaevaluation and the Australian new public management academia. In R. de la F. Ballesteros (coord.) *Yesterday and today of international education* (pp. 93-146). Madrid: Wisteria.
- Massaro, V. (2013). TEQSA and the holy grail of outcomes-based quality assessment. In S. Marginson (ed) *Tertiary Education Policy in Australia*, (pp. 49-58). Melbourne, Victoria: Centre for the Study of Higher Education.
- Mathews, T. M. & Hansen, C. M. (2004). Ongoing assessment of a university foreign language program. *Foreign Language Annals* 37(4): 630-640.

- McAlpine, D. and Dhonau, S. (2007). Creating a culture for the preparation of an ACTFL/NCATE program review. *Foreign Language Annals* 40(2): 247-259.
- McCluskey, G., Lloyd, G., Kane, J., Riddell, S., Stead, J. and Weedon, E. (2008). Can restorative practices in schools make a difference? *Educational Review* 60 (4): 405–417.
- McTaggart, R., D. Caulley, and S. Kemmis (1991). Evaluation traditions in Australia: Distillation of the old, wellspring of the new. *Evaluation and Program Planning* 14(3): 123-130.
- Mertens D. M. & Wilson, A. T. (2012). *Program evaluation: theory and practice. A comprehensive guide*. New York: Guilford Press.
- Meyers, D. (2012). *Australian Universities: A Portrait of Decline*. Palmwood: AUPOD
- Miron, G. (2004). Evaluation Report Checklist. Evaluation Center, Western Michigan University. <https://wmich.edu/sites/default/files/attachments/u350/2018/eval-report-miron.pdf>
- Molony, J. (2000). Australian universities today. In T. Coady (ed.) *Why Universities Matter: A Conversation about Values, Means and Directions* (pp. 72-84). Sydney: Allen and Unwin.
- Morris, M. (2006). Addressing the challenges of program evaluation: One department's experience after two years. *The Modern Language Journal* 90(4): 585-588.
- Murray, D. and B. Dollery (2006). Institutional breakdown? An exploratory taxonomy of Australian university failure. *Higher Education Policy* 19(4): 479-494.
- Nevo, D. (2009). Accountability and capacity building: cCn they live together? In K.E. Ryan and B.J. Cousins (eds.) *The Sage International Handbook of Educational Evaluation* (pp. 291-304) Los Angeles: Sage.
- Nevo, D. (2006). Evaluation in education. In I. Shaw, J.C. Greene, and M.M. Mark (eds.) *The Sage Handbook of Evaluation: Policies, Programs and Practices* (pp. 441-460). London: Sage.
- Niessen, T. J. H., Abma, T. B., Widdershoven, G. A. M., Van der Vleuten, C. P. M. (2009). Learning in (inter)action: A dialogical turn to evaluation and learning. In K.E. Ryan and B.J. Cousins (eds.) *The Sage International Handbook of Educational Evaluation* (pp. 375-395) Los Angeles: Sage.
- Nilsson, N. and Hogben, D. (1983). Metaevaluation. In House, E.R. (ed.) *Philosophy of Evaluation (New Directions in Program Evaluation, 19)* (pp. 83-97). San Francisco, CA: Jossey-Bass.
- Norris, J. M. (2006). The why (and how) of assessing student learning outcomes in college language programs. *The Modern Language Journal* 90(4): 576-583.
- Norris, J. M. (2008). *Validity Evaluation in Language Assessment*. Frankfurt am Main: Peter

Lang.

- Norris, J. M. (2009). Understanding and improving language education through program evaluation: Introduction to the special issue. *Language Teaching Research* 13(1): 7-13.
- Norris, J. M., Davis, J. McE., Sinicrope, C. and Watanabe, Y. (eds.). (2009). *Toward Useful Program Evaluation in College Foreign Language Education*. Honolulu: University of Hawai'i, National Foreign Language Resource Center.
- Norris, J. M. and Watanabe, Y. (2013). Program evaluation. In C. A. Chapelle *The Encyclopedia of Applied Linguistics* (pp. 1-6). Chichester: Wiley-Blackwell.
- Norris, N. (1998). Curriculum evaluation revisited. *Cambridge Journal of Education* 28(2): 207-220.
- Omaggio, A. C., Eddy, P. A., McKim, L. W. and Pfannkuche, A. (1979). Looking at the results. In J. K. Phillips (ed.), *Building on Experience-Building for Success* (pp. 233-270). Skokie, IL: National Textbook Co.
- Owen, J.M. (1991). An evaluation approach to training using the notion of form: An Australian example. *Evaluation Practice* 12(2): 131-139.
- Owen, J.M. (2003a). Evaluating educational programs and projects in Australia. In T. Kellaghan and D.L. Stufflebeam (eds.) *International Handbook of Educational Evaluation* (pp. 751-768). Dordrecht: Kluwer Academic.
- Owen, J.M. (2003b). Current developments in evaluation in Australia: Issues for evaluation practice. In N. Lunt, C. Davidson and K. McKegg (eds.) *Evaluating Policy and Practice: A New Zealand Reader* (pp. 240-242). Auckland: Pearson Education.
- Owen, J.M. (2004). Can accountability systems enhance program quality? *Australasian Evaluation Society International Conference, 13-15 October-Adelaide, South Australia*.
- Owen, J.M. and P.J. Rogers (1999). *Program Evaluation: Forms and Approaches*. London: Sage.
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text*. Thousand Oaks CA: Sage.
- Patton, M. Q. (2002). A vision of evaluation that strengthens democracy. *Evaluation* 8(1): 125-139.
- Patton, M. Q. (2012). *Essentials of utilization-focused evaluation*. Thousand Oaks, CA: Sage.
- Polya, G. (1957) *How to Solve It. A New Aspect of Mathematical Method*. 2nd Edition. Princeton: Princeton University Press.
- Premfors, (1989). *Policy Analysis*. Lund: Studentlitteratur.

- Preskill, H. and Torres, R. T. (1999). *Evaluative inquiry for learning in organizations*. Thousand Oaks, CA: Sage.
- QSR International (1999) *NVivo Qualitative Data Analysis Software* [Software]. Available from <https://qsrinternational.com/nvivo/nvivo-products/>
- Raizen, S.A. and P.H. Rossi (1981). *Program Evaluation in Education: When? How? To What Ends?* Washington, DC: National Academy Press.
- Rea-Dickins, P. and Germaine, K. P. (1992). *Evaluation*. Oxford: Oxford University Press.
- Reboloso, E., Fernández-Ramírez, B. & Cantón, P. (2009). Quality criteria for self-evaluation in higher education. *Journal of Multidisciplinary Evaluation* 6(11): 16-31.
- Reid, I. (2009). The contradictory managerialism of university quality assurance. *Journal of Education Policy* 24(5): 575-593.
- Review of Commonwealth Administration. 1983. *Review of Commonwealth Administration Report*. Canberra: Australian Government Publishing Service.
- Richards, J. C. (2001) *Curriculum Development in Language Teaching*. Cambridge: Cambridge UP.
- Rivera, G. & Matsuzawa, C. (2007). Multiple-language program assessment: Learners' perspectives on first and second year college second language programs and their implications for program improvement. *Foreign Language Annals* 40: 569–583.
- Ryan, K. and I. Feller (2009). Evaluation, accountability, and performance measurement in national education systems: Trends, methods, and issues. In K.E. Ryan and J.B. Cousins *The SAGE International Handbook of Educational Evaluation* (pp. 171-190). Thousand Oaks, CA: Sage.
- Sanders, R. F. (2005). Redesigning introductory Spanish: Increased enrolment, online management, cost reduction, and effects on student learning. *Foreign Language Annals* 38(4): 523-532.
- Saunders, M. (2011). Setting the Scene: The four domains of evaluative practice in higher education. In M. Saunders, P. Trowler, and V. Bamber (eds.) *Reconceptualising Evaluation in Higher Education: The Practice Turn* (pp. 1-17). Maidenhead: McGraw- Hill Education.
- Saunders, M., P. Trowler and V. Bamber (eds.) (2011). *Reconceptualising Evaluation in Higher Education: The Practice Turn*. Maidenhead: McGraw- Hill Education.
- Scarino, A. & Liddicoat, A. J. (2009). *Teaching and Learning Languages. A Guide*. <http://www.tllg.unisa.edu>.

- Scriven, M. (1969). An introduction to meta-evaluation. *Educational Product Report 2*: 36-38.
- Scriven, M. (1981) *The Logic of Evaluation*. Inverness, CA: Edgepress.
- Scriven, M. (1998). Minimalist theory: The least theory that evaluation practice requires. *American Journal of Evaluation* 19(1): 57-70.
- Scriven, M. (2007). *Key Evaluation Checklist*. <http://www.wmich.edu/evalctr/checklists>
- Scriven, M. (2009). Meta-evaluation revisited. *Journal of Multidisciplinary Evaluation* 6(11): iii-viii.
- Schwandt, T. 2009. Toward a practical theory of evidence for evaluation. In S. I. Donaldson, C. A. Christie, M. M. Mark (eds.) *What Counts as Credible Evidence in Applied Research and Evaluation Practice?* (pp. 197-212). Thousand Oaks, CA: Sage.
- Schwandt, T. A., and Burgon, H. (2006). Evaluation and the study of lived experience. In I. F. Shaw, J. C. Greene, & M. M. Mark (eds.), *The SAGE Handbook of Evaluation* (pp. 33–55). London: Sage.
- Secolsky, C. and Denison, B. (2012). *Handbook on Measurement, Assessment, and Evaluation in Higher Education*, New York: Routledge.
- Shadish, W.R., T.D. Cook and L.C. Leviston (1991). *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, CA: Sage.
- Sharp, C.A. (2003). Development of program evaluation in Australia and the Australasia Evaluation Society: The early decades. *Evaluation Journal of Australasia* 3(2): 6-16.
- Shawer, S. F. (2013). Accreditation and standards-driven program evaluation: Implications for program quality assurance and stakeholder professional development. *Quality and Quantity* 47(5): 2883-2913.
- Shawer, S. F. and Alkahtani, S. A. (2012). The relationship between program evaluation experiences and stakeholder career satisfaction. *Creative Education* 3(8): 1336-1344.
- Shawer, S. F. and Alkahtani, S. A. (2013). Career development in language education programs. *Australian Journal of Teacher Education* 38(6): 17-33.
- Shiller, R. J. (2019). *Narrative Economics: How Stories Go Viral and Drive Major Economic Events*. Princeton: Princeton UP.
- Simmons, H. (2006). Ethics in evaluation. In I. Shaw, J. C. Greene, and M. M. Mark (eds.), *The SAGE Handbook of Evaluation* (pp. 244-266). London: Sage.

- Smith, N. L. (2009). Fundamental evaluation issues in a global society. In K.E. Ryan and B.J. Cousins (eds.) *The Sage International Handbook of Educational Evaluation* (p.p. 37-51). Thousand Oaks, CA: Sage.
- SSCSW (Senate Standing Committee on Social Welfare) 1979. *Through a Glass Darkly: Evaluation in Australian Health and Welfare Services, Volume 1: The Report*. Canberra: Australian Government Publishing Service.
- Stake, R. E. (1986). *Quieting Reform: Social Science and Social Action in an Urban Youth Program*. Chicago: University of Illinois Press.
- Stake, R. E. (2005). Qualitative Case Studies. In N. K. Denzin and Y. S. Lincoln (Eds.) *The SAGE Handbook of Qualitative Research* (pp. 443-466). Thousand Oaks, CA: Sage.
- Stake, R. E. (2004). *Standards-Based and Responsive Evaluation*. Thousand Oaks, CA: Sage.
- Stake, R. E., G. Contreras and I. Arbesú (2012). Assessing the quality of a university—particularly its teaching. In C. Secolsky and D.B. Denison (eds.) *Handbook on Measurement, Assessment, and Evaluation in Higher Education* (pp. 3-14). New York: Routledge.
- Stake, R. E. and Schwandt, T. A. (2006). On discerning quality in evaluation. In I. Shaw, J. C. Greene, & M. M. Mark (eds.), *The SAGE Handbook of Evaluation* (pp. 404-418). London: Sage.
- Stufflebeam, D. L. (1975). Evaluation as a community education process. *Community Education Journal* 5(2): 7-12,19
- Stufflebeam, D. L. (1999). *Program Evaluations Metaevaluation Checklist*. Kalamazoo: Western Michigan University, Evaluation Center. Retrieved from https://wmich.edu/sites/default/files/attachments/u350/2014/program_metaeval_short.pdf
- Stufflebeam, D. L. (2001a). Evaluation models. *New Directions for Evaluation*, 89, 7-99.
- Stufflebeam, D. L. (2001b). The metaevaluation imperative. *American Journal of Evaluation* 22 (2):183-209.
- Stufflebeam, D.L. and A.J. Shinkfield (2007). *Evaluation Theory, Models and Applications*. San Francisco: Jossey Bass.
- Traveller, A. (2014). New Zealand. In A. Pausits, G. Zheng and R. T. Abede (Eds.) *New Public Management in Higher Education. International Overview and Analysis* (pp. 507-570). Edition Donau-Universität Krems.
- Trochim, W. (1998). An evaluation of Michael Scriven's "Minimalist theory: The least theory that practice requires." *American Journal of Evaluation* 19(2): 243-249.

- Trowler, P. (2011). The higher education policy context of evaluative practices. In M. Saunders, P. Trowler, and V. Bamber (eds.) *Reconceptualising Evaluation in Higher Education: The Practice Turn* (pp. 18-32). Maidenhead: McGraw- Hill.
- Turner, G. & K. Brass. 2014. *Mapping the Humanities, Arts and Social Sciences in Australia*. Canberra: Australian Academy of the Humanities
- Vidovich, L. (2002). Quality assurance in Australian higher education: Globalisation and “steering at a distance”. *Higher Education* 43(3): 391-408.
- Weiss, C. H. (1986). The stakeholder approach to evaluation: Origins and promise. In E. R. House (ed.) *New Directions in Educational Evaluation* (pp. 145-157). New York: Routledge.
- Watanabe, Y., Norris, J. M., and González-Lloret, M. (2009). Identifying and responding to evaluation needs in college foreign language programs. In J. M. Norris, J. McE. Davis, C. Sinicrope, and Y. Watanabe (eds.), *Toward Useful Program Evaluation in College Foreign Language Education* (pp. 5-56). Honolulu: University of Hawai’i, National Foreign Language Resource Center.
- Williams, R., G. de Rassenfosse, P. Jensen & S. Marginson (2013). The determinants of quality national higher education systems. *Journal of Higher Education Policy and Management*, 35(6): 599-611.
- Wingate, L. A. (2010). Meta-evaluation: Purpose, prescription, and practice. In E. Bader, E. Peterson and P. McGraw, B (Eds), *International Encyclopedia of Education*. San Diego: Elsevier.

Endnotes

¹ Owing to word length, tables including individual case studies evaluation materials are not included in the submitted thesis. This additional information will be part of the deposited on-line dissertation version.

² Stakeholders are defined as “people having a ‘stake’ or interest in the outcome of the evaluation” (Mackay 1988, p. 35), hence, the need to include those most affected by an evaluation into the evaluative process to raise the likelihood of evaluation use, and to balance issues of power between different evaluation players (Weiss 1986). In fact, the stakeholder approach to evaluation was created to “redress the inequitable distribution of influence and authority” (Weiss 1986, p. 195).

³ However, Dahler-Larsen (2009) warns that “if process use becomes the main criteria of success, it is almost always possible to identify some positive side of what was called an evaluation process. In the extreme and perhaps caricatured situation, there will be more internal satisfaction with the evaluation process the less productive it is.” (p. 314).

⁴ This workshop was organized by the National Foreign Language Resource Center at the University of Hawai’i in 2007.

⁵ For a full account of each of the seven evaluation cases see Norris, Davis, Sinicrope, and Watanabe 2009.

⁶ Birckbichler (2006) comments on the two other uses that Aiken et al 1979 differentiate, apart from instrumental (direct action), which are the conceptual use (to affect the thinking of policy makers) and the symbolical use (advocacy or criticism).

⁷ Kemmis (1986) is one of the first evaluation theorists in Australia to elaborate on how metaevaluation may play a role in the evaluations of educational programs. He conceived metaevaluation not so much as an oversight mechanism to warrant a quality evaluation process but as one with a similar role to that of evaluation itself, that is, to expand the existing critical debate around a program. The need for metaevaluation, Kemmis states, may originate from the sponsors or from the program practitioners themselves. Sponsors, on the one hand, as providers of resources, may wish to ensure that they are adequately allocated, and as co-participants of a program, that the evaluation is feeding into the critical debate about the program. Sponsors may also use metaevaluation as a way of monitoring the degree of responsiveness of the primary evaluation to the needs of the program. On the other hand, program practitioners may also benefit from a metaevaluation in terms of ensuring that the plurality of perspectives, concerns and interests in the program community are incorporated into the debate.

⁸ Shiller (2017), winner of the Nobel Prize, in his book *Narrative Economics* studies “the spread and dynamics of popular narratives, the stories, particularly those of human interest and emotion, and how these change through time, to understand economic fluctuations.” Shiller claims that when we want to understand significant events “we rarely focus on the important narratives that accompanied them” (p. 5). A narrative is defined as “a simple story or easily expressed explanation of events that many people want to bring up in conversation or on news or social media because it can be used to stimulate the concerns or emotions of others, and/or because it appears to advance self-interest.” (p. 4). Therefore, in order to understand the views believed by people in a particular era it is necessary to immerse ourselves in the existing narratives at the time. Shiller also considers the unknown ability of some narratives to spread while others don’t. However, he argues that “While we may never be able to explain why some narratives “go viral” and significantly influence thinking while other narratives do not, we would be wise to add some analysis of what people are talking about [...]” (p. 10).

⁹ In New Zealand, the increased regulatory function of the State in the Tertiary sector is driven by the goal of using HE institutions to build a stronger economy. The State controls by requiring HEIs to produce a strategic plan that aligns with the national Tertiary Education Strategy (TES). However, an OECD report signaled the risk of becoming “compliance” rather than achieving goals due to the tendency to overemphasize details and compliance (Traveller 2014). Though not Australia per se, this information may be relevant “since Halligan (2011) states that “In addition to specific reforms in Britain (for example privatization and executive agencies) individual country programs gained international significance with New Zealand’s ‘public management model’ being influential (Boston et al. 1996)” (p. 84).” Thus, universities in Australia may have become a means to meet economic goals set up by government. In this context, one of the potential purposes of institutional evaluation, thus, could be to control strategies are put in place to meet such goals.

¹⁰ Confronted with a refusal by the University of Melbourne Press, which raised concerns in as well as beyond the university campus, this book was finally published by the independent Australian publisher Allen and Unwin, today out of print, but the arguments and process of refusal for its publication, as the authors recount in the book, present similar disconcerting characteristics to Meyers’ more recent portrayal. It is striking that

the voices raising these types of concerns are still taking place today as 19 years ago when Coady's book was published.

¹¹ Harman (1998) compares the style used in Australian universities before the 1980s to the British approach to quality assurance. On the one hand, at institutional level it consisted of periodic departmental and course reviews and relied on the help of teaching and learning centres for improving the quality of teaching, and on the other, the system was under the control of Commonwealth and State committees operated by professional associations. Therefore, the contrast highlights current quality assurance processes driven by economic needs and rankings, in alignment with governmental policy requirements, thus, deeming educational quality as a secondary matter.

¹² In addition, and in support of the main research questions, these additional self-awareness queries contributed to inform the metaevaluation:

1. To what degree did teachers participate in, or were involved in the evaluation so that they gain a more comprehensive understanding of their programs?
2. Were a variety of data collection methods used to ensure a multiple perspective when interpreting?
3. Were different data sources and methods used to attain triangulation thus helping to contextualize findings and in turn staving off inaccurate and politically motivated interpretations?
4. What reporting formats were used to communicate findings? To what audiences were findings reported? Were different styles of reporting used to communicate findings that conveyed an accurate understanding of the programs so that likelihood of evaluation utilization is increased and awareness about the value of PE is raised?
5. What role did the evaluator(s) play?
 - 5.1. Was the evaluator a traditional 'expert assessor' or did the evaluator adopt an educative orientation by helping language educators understand the potential contributions to be made by the evaluation or both?
 - 5.2. Did the evaluator engage in considerable learning about the language program, its learners, teachers, social circumstances etc.?
 - 5.3. Was the evaluator involved proactively in the evaluation design and the implementation of methods to respond to intended uses and users?
 - 5.4. Did the evaluators construct their efforts within frameworks that made sense to the stakeholders and audiences related to language programs?
 - 5.5. Did the evaluator pursue a deep and meaningful relationship with the specific language teaching milieu?

¹³ *Confidence* is an underscored value according to Reboloso et al (2009). While not considered traditionally as a criterion for scientific inquiry, it is "what brings us closer to the issues of responsiveness and utility – characteristics of applied research, social intervention and management" (p. 28).

¹⁴ For a full account of these issues refer to Martínez Marco 2013.

¹⁵ Regarding causal analysis, the fourth inquiry mode suggested in the ERE framework, it is left out due to lack of sufficient data.

¹⁶ Theorists and practitioners of metaevaluation have found both restrictions in the data (Cooksy and Caracelli 2005) and in accessing certain types of data. To overcome them, metaevaluators are advised to "achieve buy-in from the involved parties" which may entail various attempts (Wingate 2010, p. 772).

¹⁷ These universities are considered Australia's leading research-intensive: <https://go8.edu.au/about/the-go8>

¹⁸ The findings have been published and are available at <http://www.lcnau.org/proceedings/2013-proceedings/>

¹⁹ Proof of endorsement of project can be found at to <http://www.lcnau.org/projects/>

²⁰ Evaluation, furthermore, tends to be conflated with testing (Jacobson 1982) and with instructor evaluation (Jarvis and Adams 1979) and according to Mackay (1988) the lack of interest in program evaluation in language education is due also to a matter of priorities whereby immediate program needs such as planning and implementation are set in motion while leaving evaluation unattended. Jacobson (1982) refers also to practical matters such as the lack of professional evaluation expertise in language programs to justify that, as Omaggio (1979) states, "the foreign language community has been strangely quiet about the problem, which leads one to suspect that evaluations are not being carried out with any regularity or in any systematic fashion." (p. 240).

²¹ The initials shown in the Figure are used as a code to anonymously identify each of the participants in the interviews.

²² See online Appendices for all interview questions.

²³ In Stufflebeam and Shinkfield (2007) it is stated that resources allocated for metaevaluations represent 2% of the total primary evaluation budget (Wingates 2010).

²⁴ A copy of this checklist is provided in the online Appendices

²⁵ The rigor of repeated readings of data, of field notes and the refinement of the code book as a result of the summary of codes is particularly emphasized as an analytical step since it adds validity to the analysis (Alkin 2011).

²⁶ This call to the Review Panel is made one more time at the end of the Introduction chapter expressing that “We trust that the Review Panel will [...]” take into consideration the analysis, results and recommendations “in this paper” and their hope that the Panel’s decisions “will be fair ones based on data, sound analysis, and coherent policy”. Therefore, this evidence suggests the high stakes nature of this review, and its repetition, an indication of CS6’s self-reviewers’ potential lack of confidence in the fairness of the primary review process.

²⁷ In CS4’s ES, particularly, the summary of challenges reveals a grammatical irregularity that can be connected with an important finding brought up in the verbal data: After the “first”, “second”, “third” and “fourth” challenges are presented and explained, the fifth one is introduced with “Finally” thus implying it is the last challenge to be presented. However, what follows are two further clauses introduced by “Moreover” and “Similarly” which raise two further challenges completely unrelated to the fifth one. Interestingly, the focus of challenges (6) and (7) on the lack of explicit acknowledgement by the University about the importance of language study in meeting the University Strategic research and teaching plans, and on the marketing literature not emphasizing enough the “cultural, educational and professional benefits” of language study at the tertiary level respectively, are issues raised in a recommendation commented by interviewee (C), who was a member of the Review Panel in this review:

C67. [...] We said that it would be very helpful to the health of languages at the (U3) if, if there were to be in the university’s statements about its purposes some mention of languages as being useful, significant, important [at the university].

and which interviewee (C) continues “was excised from the final report”. This finding, therefore, clearly indicates in correspondence with the previous finding that the last two challenges, that is, (6) and (7), may in fact not have been intended to be included in the first place. But, the question is who decides what is and what is not to be included in the “final report” and who writes this report if it is not the reviewers themselves?

²⁸ This same descriptive paragraph about the program appears in the Conclusions of the report, thus, emphasizing the alignment of program with Strategic Plans not only within the university but at a national level.

²⁹ In fact, the categorization of ‘self-reviewers’ for CS6 is the clear identification in the TP of the authors of the report with the authors of the review.

³⁰ In CS4 and CS5, these areas are clearly connected to the ones presented in a section named “Terms of Reference” (ToR) of the two reports. As a full ToR section is also included in CS2, these are analyzed in detail as an additional metaevaluation segment later on on this chapter.

³¹ The identification of these areas in CS2 and CS3 denotes a review within another review, whereby a review report of the evaluand is submitted to a panel of evaluators, exactly as in CS6.

³² Since in NPM universities the aim is to become independent from government funds, attracting high fee paying postgraduate international students thus increasing number of MA and postgraduate courses is an established strategy (Martínez Marco and Pérez de León 2019)

³³ The breaking down of these plans in CS4 is not indicated in the ToR but in “Appendix C”.

³⁴ Grammatical error in report.

³⁵ Notice how the self-reviewers alert against a problematic initiation of the review by using a term with a negative connotation (e.g. allegations) as well as by informing about extra-work data beyond the questions in the Review Notice.

³⁶ A question may be raised here about the extent to which these ToR were designed for CS6 alone or were the same ones which the Review Panel to whom CS6’s report was to be submitted may have received.

³⁷ A full explanation of the recommendations in these case studies is exposed under the Recommendations section of the ERC at the end of this chapter.

³⁸ Why, it may be asked, this persistent need to highlight so emphatically the staff’s positive engagement with a review?

³⁹ This is a relevant metaevaluation finding since in the pilot study carried out before this thesis, data showed that the Australian universities where interviewees worked did not have cycles of reviews. Therefore, a larger metaevaluation study was needed to show that cycles of reviews do take place though not in all Australian universities.

⁴⁰ Memorandum of Understanding.

⁴¹ This outcome can be confirmed in CS5’s report since it appears as one of the recommendations.

⁴² The analysis of data in CS5 report and descriptions in the verbal data from both (M) and (I) lead to identify interviewee (M) as one of the panel members in CS5, as is clarified later on.

⁴³ It must be noted how (M) comes into the review as the rest of the reviewers with a preconception about the program. In fact, it seemed to have been picked up by (I) during the review and considered an important obstacle in the review

I19. [...] that we were weak disciplinary, disciplinary weak, which was the preconception of this woman anyway.

I34. [...] I mean, this ["we were not er sort of enough disciplinary trained"] was a big issue, [...] whereas for (M), the conflict came about due to the commissioners' concern and "therefore colleagues in the languages had a reason to be very defensive". In fact, (M)'s preconception can be reaffirmed through an experience outside the review in connection with the marking of students work at CS5's university suggesting incompetence amongst the languages staff:

M22. [...] they've got a way of marking, of correcting students work that is utterly inappropriate, yeah, I think. [...] and so they're the kind of people who think that if their marking is particularly stringent that means they're increasing-they've got high standards.

M24. They don't] actually understand that the level comes from somewhere else, doesn't come from there.

⁴⁴ According to an online search.

⁴⁵ The question is where could staff's persistent resistance to the second review's findings stem from? At any rate, this metaevaluation finding suggests that in order to anticipate the level of commitment to recommendations ensuing from evaluation results would be to conduct a needs analysis to gather staff's perceptions before planning and conducting a review in order to avoid the waste of time, effort and resources and tackle the issue from a different angle.

⁴⁶ The commissioning of reviews by university authorities can be attributed to the current "self-managing and self-improving" characteristic of NPM hierarchies within universities with one of the various tasks they are responsible for being "evaluation and reviews" (Traveller 2014).

⁴⁷ <https://en.oxforddictionaries.com>

⁴⁸ This finding thus raises the question about what in this context may motivate potential reviewers getting involved with such painstaking unpaid work.

⁴⁹ An interesting coincidence can be identified in the composition of the RP in CS4 and CS5. On the one hand, the internal member in CS5's RP seems to be the external member with the "eclectic background" in CS4. On the other hand, one member in CS4's RP share common academic link to one in CS5. Does this mean that there is a tendency to invite reviewers of this kind of rank or even same individuals across universities?

⁵⁰ The action of covering names of participants in CS5 is supposedly done for the purpose of securing confidentiality which means that the provider of this report did not trust the confidential nature of this metaevaluation research study.

⁵¹ An online search also allows to confirm that (M) was the member of the RP holding such an award at the time.

⁵² The metaevaluation findings about (M) so far exposed help identify him as the "not-adequate" member of CS5's RP.

⁵³ In the verbal data, (A), who identifies himself as the Head of the program at the time, states that in fact, the report and most of the work in CS6's report is done by him:

A266. Sure]. Yeah. This one, I made a 2 version. Everything I wrote is my work and anything, other than financial information—

Q267. You were the convener of the program at that time.

A268. Yeah. I wrote the whole thing so um there are a few things like er— [...]

⁵⁴ Therefore, this evidence shows that while the primary review report to which CS6 is embedded is not made available, CS2 report in which CS3 is embedded is made available. The question to be asked is what criteria is used in order to provide one and not the other.

⁵⁵ Metaevaluation research on this potential finding has been confirmed through an online search where the university exposes standard "Terms of Reference" for a number of "Committees" for various operations, one of them being "Academic Policy and Program Review".

⁵⁶ Verbal data suggests, in fact, that finances is the principle that guides other university strategic operations such as the hire of managers, also suggested by (I) as seen above, who in turn use the same principle to guide reviews. In this case, while (A) illustrates this idea through the example of a past review four years earlier, this seems to be perceived as a widespread practice:

Q85. Did she er commission it?

A86. Yeah, she brought in some external people. Er I think she already had agenda so, you know, when dean was brought in probably already had a mission of basically save money. So saving money might mean flexible delivery and reduced face to face hours um reduction of (xxx) numbers, offerings and cut some staff, and make the budget better. I think there was a—

Q87. The main drive [for—

A88. Yeah,] absolutely. It's all financial.

and which is heavily criticized by (A) to the extent of “complete despair” because this focus seems to be accompanied by a lack of interest in students (“Students, students’ experience was the least of their concern.”), or driven simply by prioritizing finances over other university important considerations:

A352. Oh, well, that’s what I wanna know. [...] because they say like um they talk about, you know, internationalizing students as a strategy plan and they talked even about exchange and we were doing exactly that and they just kind of think ‘oh, yeah, that’s important but not as important’ and just kind of take strategy, vision, policy and I used that as a guiding line: Teaching content I— there are several things I believe students should learn but as a guiding principle of the organization, the program within the organization, I was using that as a guiding principle and they just said ‘eh, okay’, so I mean, what can we do?

⁵⁷ It must be noted that CS3’s description is an almost identical copy of the one found in CS2 which raises the question of authorship. Could this be evidence of identical authorship in CS2 and CS3? Otherwise, who copied from whom? While answers to these questions are not available, this finding does reveal an internal and administrative orientation of the review.

⁵⁸ Although no evidence is provided to support this claim, it does clearly imply that the author of this report expresses an inclination towards supporting this efficiency trend.

⁵⁹ Here “language courses” must be understood as the staff responsible for the teaching of language courses.

⁶⁰ Aboriginal and Torres Strait Islander

⁶¹ Non-English-Speaking Background

⁶² Culturally and Linguistically Diverse

⁶³ Higher Degree Research

⁶⁴ In CS6, a full citation of references is provided in the footnotes systematically.

⁶⁵ CS2/CS3 are an exception since they are triggered by an “inaugural” cycle.

⁶⁶ Though a speculative inference, this evidence could lead to identify this individual as someone with a role similar to the “secretariat” in CS5 included in the Membership of the Review Panel section and therefore, as the person behind the writing of this summary.

⁶⁷ According to the ERC, this is not a checkpoint to be included in the Methods chapter, and since the identification of reviewers is already reflected in the TP, this information is redundant unless it is considered a relevant issue to emphasize in this review process.

⁶⁸ As will be seen in the metaevaluation findings under Appendices later on, only two (2) of the submissions are included.

⁶⁹ Is this role to be read as commercial? If so, it may signal, on the one hand, a commercial trend in this university, which in this case has positive connotations given its uniqueness. In fact, as can be inferred from evidence in the next metaevaluation chapter, this denomination may have been assigned to this individual in recognition of the contribution of this project to the School, which is praised particularly for its capability to contribute to the sustainability of the School in both CS2 and CS3 repeatedly.

⁷⁰ The quote below from (I)’s verbal data reveals additionally that the focus in the School’s submission is quantitative data, particularly, outcomes in relation to the administrative side of the program with an emphasis on student numbers:

I19. Well, I have produced some of the data, not all the data that was consider was er I mean, enrolment statistics and suppose er I mean, exchange, I also produced— because, I mean, I know how to handle data. I was a mathematician in a previous life and so er I also show with the data, that I gave the head of school who gave it to them, [...]

⁷¹ It is noticeable how despite the similarity of this section with the contents in CS2 and CS4’s Appendices, in CS5, it is a main section in the report.

⁷² It is interesting to note how the title “Staff attending the review” reveals a conception of the role of staff not as participants but as attendees, which suggests a limited understanding of participation as ‘meetings with review panels’. Moreover, does this concept imply that participants are ‘free to attend’ or are they compelled to attend by policy?

⁷³ The Appendices from this case study have not been made available for the purpose of this thesis.

⁷⁴ The contradictory conclusions in consecutive review reports that this quote reveals leads to claim that, at this university, past reviews are not taken into account when reviewing a program and therefore that ongoing developmental work is not usually considered in reviews or if it is, there is no follow-up.

⁷⁵ The last Appendix containing student letters has not been made available for the purpose of this research.

⁷⁶ It must be noted that according to (A), the research work that is put into the self-review began well before the review was commissioned. The meeting with the Dean announcing the closure of the program and the fact that Dean was a new replacement sparked an interest in (A) to write about the program particularly to ensure the new dean understood correctly the programs' true status:

A340. Well, when that happened, I immediately start writing about the (L6) program um before this review happens. I wrote— I was—

A344. Yeah, but in between I thought 'I should write up how wonderful (L6) program do because these idiots probably doesn't know' and then I still had faith in the system. I thought 'he's new, he's out of his depth, he doesn't know, so if I explain what contribution (L6) program is making— financially I didn't have the information but er like what students did, the um— were coming those (D3) (L11) language course students, and er contribution to the er exchange program. (L6) is the second biggest destination in the (U4) as an exchange destination, and second to (COU5), right? And er (CO1) area exchange was actually nonexistent without (L6) program, it was 90% or 96% was (L6) program in (AL) area.

⁷⁷ Equivalent Full Time Study Load

⁷⁸ Course Evaluation Questionnaires

⁷⁹ Tertiary Entrance Rank

⁸⁰ Faculty of Education.

⁸¹ Graduate Course Evaluation Questionnaire.

⁸² While the focus of the review anticipates the search for negative reasons that can lead students to discontinue the program, positive and negative findings can be clearly discerned.

⁸³ Metaevaluation anonymity code.

⁸⁴ Metaevaluation anonymity code.

⁸⁵ Metaevaluation anonymity code.

⁸⁶ According to <https://en.oxforddictionaries.com/>, 'lean' is defined as "(of an industry or company) efficient and with no wastage".

⁸⁷ The curricular reviews, according to the report, produced a report discussed between chairs of programs and departments and a Review Panel without specifying whether this panel was internal or external. If this is not a formal review, what makes a review 'formal'?

⁸⁸ Another example of targeted hiring for making the program successful, as seen later, associated with increasing enrolments.

⁸⁹ This metaevaluation finding, nevertheless, cannot be considered reliable since these appendices have not been made available for the purpose of this research.

⁹⁰ Even though CS6 is framed within a Quality Assurance cycle of reviews, the only reference found to work from past reviews is the marketing plan.

⁹¹ English as a Second Language.

⁹² According to <http://elearningincubator.com.au/>, this program is a "cloud-based e-learning start-up solution" that helps organizations such as "smaller universities" to deal with "the cost of purchasing and maintaining hardware, software and the expertise to go with it" which "can be a huge barrier to even the smallest venture into technology assisted learning." In other words, it is a way of making e-learning more cost efficient, in line with the UEM as described in CS3.

⁹³ The financially robust and the one yet to prove financially viable are in fact the two interdisciplinary programs illustrated earlier (see p. 74).

⁹⁴ It must be pointed out that it was the justifications found in CS2, CS4 and CS5 that made this section in each of the case studies appear initially as the Results chapter and were treated as such during analysis. Further work evidenced their summary nature though in the end they can be qualified as a hybrid between a Results Chapter and a Summary.

⁹⁵ In-text error.

⁹⁶ It must be noted that the anonymous writer's subjective stand in CS4 is reflected not only in this chapter but also in other areas of the report. For example, in the ES, after highlighting the willingness of the teaching staff to engage in "opportunities" which the "review panel felt", the writer shows an enthusiastic agreement with such finding: "Indeed, there appear to be some very realistic opportunities for growth [...]; "Similarly there are numerous opportunities for [...]" which seem to be connected with increasing enrolments.

⁹⁷ However, the question to ask is why is this recommendation excised when it did not imply the use of “extra-resources” which was the only restriction stated in the ToR? Since no explanation is provided, it may be claimed that restrictions, whether explicit or implicit, are at the discretion of commissioners which interviewee (C) interprets as authoritarian behavior: “we were not there to give the vice chancellor advice on what his statements and policies should be” which, as in this case, may end up in the removal of unwanted information. Furthermore, in connection with a previous metaevaluation finding (see footnote 7) based on a grammatical incongruity in the text, it may be argued that the forced affixing of challenges 5 and 6 in the text in the ES clearly responds to a last-minute decision to include these after the removal of the excised recommendation.

⁹⁸ The concept of ‘final report’ in this context needs to be associated with the re-writing of the Reviewers report (bureaucratic version following NPM principles)

⁹⁹ See p. 26-27.

¹⁰⁰ It must be noted that should CS2’s findings be considered the Results chapter, as the anonymous writer suggests in the Introduction, evidence to support findings is still scarce. For example, in connection with the reputation mentioned above, under “School Positioning and Strategy” this finding is introduced by praising the reputation of the School as “a centre of excellence” with no evidence to prove it. Likewise, under “Curriculum”, “diversity’ is presented as the “major contributor” for the School’s “positive reputation” “both in Australia and internationally”. However, rather than showing manifestations of such reputation, “diversity” is explained at length and largely in terms of the School’s capability to come up with solutions that not only don’t drain resources but generate income as well as in response to University “imperatives”, thus revealing a demonstration of School’s success in complying with priorities.

¹⁰¹ This is a clear example of how reviews, or at least one of the purposes of reviews, are used to improve the financial viability of the program.

¹⁰² A question that is raised in this section is what those relevant departments are and why those countries need to be “consistently included”.

¹⁰³ The following section on “The Bachelor of Languages” in CS4, justifies this program explaining how this degree may help to attract particularly “international students”.

¹⁰⁴ While interviewee (C) clarifies the arguments supporting this selection, that is, due to the tradition of background speakers in the area, according to the discussion, the reasoning seems to be in connection with helping meet “future demand for courses” and the University “to capitalise on its market advantage” thus increasing its competitive capability.

¹⁰⁵ Notice how the review questions in CS6, rather than being research-oriented, demand that the program creates plans for actions to deal with the issues mentioned in each question.

¹⁰⁶ However, it should be asked whether these questions were in fact formulated by the Review Panel itself or the commissioners since the university has chosen not to make available the report from the primary review for the purpose of this research, or perhaps whether the self-reviewers perceive or assume that the commissioners and Review Panel represent the same group of stakeholders.

¹⁰⁷ This suggestion leads to understand the writer’s expression of his pride in serving the University in achieving priority goals and at the same time a critical, skeptical or even cynical view of the level of ‘servitude’ that the compliance required by such achievement involves.

¹⁰⁸ The identity of such individuals is unknown as well as the process by which it has been carried out.

¹⁰⁹ Should “School Submission Review” be read as a ‘Self-review’? The answer to this question cannot be answered since this “submission” has not been made part of the Appendices, as mentioned above. However, when interviewee (C) is asked about the level of engagement of program staff in the review, the answer seems to suggest that it was:

C75. [...] it turned out that they [“the people who we interviewed”] were the people who’d put the figures together.

¹¹⁰ This finding reveals the institutional character of the report.

¹¹¹ It is interesting to note how this comment implies on the one hand the internal character of the memo letter and the plan. In fact, access to this review report and plan was unsuccessful after requesting it to upper management and only made available thanks to an outside intervention. On the other hand, it also reveals the practice of keeping names unknown behind the “Faculty” and the “University” labels. But, how can then responsible individuals be held accountable when no people’s names are identified?

¹¹² The purpose of these key messages is to ensure that the program tackles 6 main topics when “communicating about” the program which include specific actions in order to make the program competitive.

¹¹³ Under the theoretical underpinnings of NPM in the New Zealand HE model, particularly from *Public Choice Theory*, universities are funded according to a performance-oriented Investment Plan where universities compete for funds and are assessed according to research degree completion, types of programs or courses

offered, external research income, number of valid enrolments and quality evaluation (Traveller 2014). In this model thus it is easy for Government to use whole-university evaluation reports to control compliance of prioritized goals. Under such pressure, the role and characterization of evaluation processes for quality assurance within universities, may become tools for prioritizing and providing funds for particular departments, schools, etc. according to compliance of specific goals which clearly confers evaluation the role of policy instrument.

¹¹⁴ As an example, see this media article: <http://www.couriermail.com.au/news/queensland/staff-at-the-university-of-queensland-are-working-in-a-culture-of-fear-with-little-trust-in-senior-management/story-e6freoof-1226609157943>

¹¹⁵ According to interviewee (A), the commissioning of the current review following the decision to close the program created further stress to the staff and suggests unprofessional practices:

A230. Active, and so they said 'oh, okay, we will think about for— this semester we'll open (L6) and we'll think about it' and 20 of March, the terms of reference is this.

¹¹⁶ This announcement is perceived by (C), who collaborated to help in this review, as an unfair and cowardly top-down decision since it implies the non-consultation of program staff:

C97. [...] Those decisions er were made without consultation of the people concerned and you have seen from the (U13) the, the er you know, it's a big knife coming down from outside with absolutely no consideration of, of— and I think that's in a way it's cowardly, it's outrageous, it's, it's er—

and the replacement of this decision "by a commitment to a process of consultation and review" is clearly viewed by all three interviewees who contributed information about CS6, as a political instrument to implement the predetermined decision:

K62. There's another. Another one that er another I think that we did last year. We helped er certainly er the closure of the university, a projective, an attemptive ((attempted)) closure of a language program. [...] There ((It)) was a review conducted by the faculty, with the aim of potentially closing a language program, and we gave, I, me personally, gave a lot of attention to helping the staff prepare a response to the review and then, going through the response, etc. etc.

C7. [...] they were trying to cut out the (L6) program [...]

A264. So the whole— my position was whole review was bollocks.

¹¹⁷ A question that may be raised is what measure is used to know exactly how much alignment is desirable or is not enough?

¹¹⁸ According to an online search.

¹¹⁹ The question is where could staff's persistent resistance to the second review's findings stem from?

At any rate, this metaevaluation finding suggests that in order to anticipate the level of commitment to recommendations ensuing from evaluation results would be to conduct a needs analysis to gather staff's perceptions before planning and conducting a review in order to avoid the waste of time, effort and resources and tackle the issue from a different angle.

¹²⁰ For a full account of each of the seven evaluation cases see Norris, Davis, Sinicrope, and Watanabe 2009.

Appendices

Appendix 1. Interview Questions

Introduction

I would like to know about your personal experience in reviews in language programs of university departments or schools.

1. How many have you been involved in?
2. Can you think of some examples that perhaps could be representative of the present situation of LPE in higher education in Australia?

Questions in reference to specific review experiences by the interviewee

3. What was the impetus or purpose of the review?
4. Who commissioned it?
5. What role did the evaluators play?
 - 5a. What process was used to select the evaluators?
 - 5b. Did they have different backgrounds? What kinds of expertise did they have? How were these reflected in the results of the evaluation?
6. What level of time, effort and resources were allocated to conduct the review?
7. What strategies or means were used in order to communicate the findings?
8. When conducting the review, to what extent was language assessment carried out at the curricular level?
9. What were the attitudes toward the review among the members of the program (teachers, administrators, etc.)?
10. What was the level of commitment of the language academic staff to conducting the review?
11. What have you learned from your participation in these review processes?
12. What has been your most positive experience?
13. What has been your most difficult experience?

General questions about Language Program Evaluation in higher education

14. How important is it that the participants are LPE literate, i.e. the stakeholders are made aware of the benefits of LPE prior to conducting the evaluation?

15. How important is it that language academic staff participate in the process of planning, designing and implementing LPE?

16. What is the relationship between the quality level of language teaching and LPE?

17. To what extent should the university language educators incorporate LPE as a component of their language program?

18. How does the level of commitment to the teaching of languages affect LPE?

General questions about Language Program Evaluation in Australia

19. I have learnt from my research that in Australia

- Not all universities use LPE on a regular basis. What is preventing language departments from using LPE more frequently? What would be needed in your opinion to make LPE a systematic practice in university language departments?
- There is scarce research conducted on LPE in Australian universities. Why do you think there is this research gap?
- Language review reports are difficult to obtain and highly confidential. What do you think this is due to?

20. How can LPE support language education? How can LPE benefit the language profession as a whole here in Australia?

21. In what ways can LPE influence language education policy in Australia?

22. Do the funding agencies have any knowledge of how language programs work?

23. To what extent do funders make a connection between reviews of language programs and language program quality?

24. Do you have any further to say about the role, function or character of language program reviews in Australian universities?

Appendix 2. Conventions for the transcription of interviews' verbal data

Format	Meaning	Example
Hyphen	Sharp cut-off of prior sound or word	undergrad-
Dash	Incomplete or interrupted utterance	No, up until—
Question mark	Voice inflection for asking a question	In Melbourne?
Exclamation mark	Voice inflection for emphatic speech	even if you think something like that, you don't say that to your staff in a faculty forum!
Underlining	Word or phrase stressed by speaker	the entire university Australia still
Brackets	Overlapping talk	1. Yes, and that is something that is commented in the Perspectives issue of 2006 [er I think. 2. Right, exactly.]
Single quotation mark	Reported speech	they just rejected it, 'We don't want this'.
Parenthesis	Possible interpretation of unclear words	so you got a (googleous) of people
Double parenthesis	Non-verbal action/researcher's comments	((laugh)) ((sigh))
Parenthesis with "xxx"	Incomprehensible words or phrases	(xxx)
Comma	Selected group of words to facilitate understanding and avoid ambiguity	administration and the departments will typically bend over backwards to make that happen, including dragging people in at 8 o'clock in the morning, on Saturday morning, if that's what you need.
Capitalization and full stop	Beginning of an utterance and its completion	So the impetus, when it came to me, it was like that.
Italics	Non-English words, expressions (excluding people's names and geographical names)/ imitation talk	<i>consapevolenza</i> <i>bluhdee bluh</i> <i>de facto</i> <i>ab initio</i>
Colon		I think the answer to that is: It would be patchy.

Appendix 3. Letters to Institutions



L. Alicia Martinez Marco
School of Languages and Linguistics
University of Melbourne, VIC 3010

Tuesday, 12 November 2013

Dear Professor X,

My name is Alicia Martinez and I am currently conducting research for a PhD in the School of Languages and Linguistics at the University of Melbourne under the supervision of Professor Tim McNamara. My project is entitled “Meta-evaluation of program evaluations in Australian university languages departmental reviews” (see www.lcnau.org for a full description) and is set within the field of Program Evaluation. It aims at using meta-evaluation (an evaluation of the character and quality of evaluation reports) as a research approach to building knowledge of evaluation practices in languages programs in Australian higher education. Through evaluating reports of program and departmental reviews from a number of Australian university languages programs, this project will shed light on the current state of language program evaluation practice in Australian universities and the findings will be used to propose recommendations to enhance evaluations processes in this domain so that it informs good practice. The project has the official endorsement of the recently created *Languages and Cultures Network for Australian Universities* (www.lcnau.org). This organization connects language academics and experts from Australian tertiary institutions and from other educational contexts with the goal of strengthening and enhancing the academic value of the study of languages and the quality of second language teaching.

My purpose in writing this letter is to request your assistance in the conduct of this research by granting me access to relevant program review reports from your institution. As you will appreciate, access to such reports is necessary for the conduct of this meta-evaluation. Given the confidential nature of this type of document, all the names of those participating in the evaluation as well as the name of your institution will of course remain completely anonymous. By allowing me to study languages program review reports, your institution would be supporting LCNAU’s work to advance the languages profession in Australia. At the same time, you would be joining other Australian universities which have already agreed to allow access to reports.

If you have any queries or concerns, please contact myself (luciamm@unimelb.edu.au) or Professor McNamara (tmcna@unimelb.edu.au). I look forward to your response.

Yours sincerely,

L. Alicia Martinez Marco
PhD Graduate Researcher

1st September 2014

Professor
University

Dear Vice-Chancellor

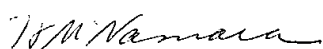
I am writing to formally request a copy of any report over the last decade that may have been made as the outcome of a review or evaluation of language programs or language departments at your university. The report would constitute part of the data for a PhD research project being carried out by a student under my supervision, Ms. Alicia Martinez Marco. The topic of her PhD, in which she is in the second year of full-time study, is 'Metaevaluation of program evaluations in Australian university languages departmental reviews'. The project aims at using meta-evaluation (an evaluation of the character and quality of evaluation reports) as a research approach to building knowledge of evaluation practices in languages programs in Australian higher education. Through evaluating reports of program and departmental reviews from a number of Australian university languages programs, this project will shed light on the current state of language program evaluation practice in Australian universities and the findings will be used to propose recommendations to enhance evaluations processes in this domain so that it informs good practice. The project has the official endorsement of the recently created *Languages and Cultures Network for Australian Universities*. Further details of the project can be found at its website, <http://www.lcnau.org/projects/>.

Given the confidential nature and occasional sensitivity of such reports, the name of the institution and the names of anyone participating in the evaluation will of course remain completely anonymous in the thesis and any resulting public presentations or publications, and as supervisor I will pay particular attention to this point.

The focus of the research is not on individual institutions or individual practices, but the manner in which such reviews are conducted and their findings disseminated in general in Australian universities, with the aim of informing best practice in future. Ms. Martinez Marco is a mature researcher with many years of experience of university language teaching (Spanish) in the United States and more recently Australia.

I hope you will be able to assist her research by giving us access to any relevant report from your University, and I look forward to hearing from you.

Yours sincerely



Tim McNamara
Professor and Deputy Head
School of Languages and Linguistics

Appendix 4. Case Studies EFI Tables

CASE STUDY 1 (Go8): Evaluation Fundamental Issues		
	Explicit	Implicit
Trigger(s)	<ul style="list-style-type: none"> • Follow-up study (qualitative following a past quantitative study) 	<ul style="list-style-type: none"> • Initiation connected to results showing high rate of discontinuation from previous quantitative study • Enhance evaluation utilization due to potential negative reception of findings from the quantitative report which may have prevented any commitment to actions
Evaluation Criteria		<ul style="list-style-type: none"> • Reviewer's professional perspective (Applied Linguistics), acknowledged limitation (no data on discontinuation available from other degrees)
Commissioner(s)	<ul style="list-style-type: none"> • Head of School where program lives 	<ul style="list-style-type: none"> •
Evaluand	<ul style="list-style-type: none"> • A language program (diploma) 	<ul style="list-style-type: none"> •
Purpose(s)	<ul style="list-style-type: none"> • To understand high rate of student discontinuations • To get an overall picture of program • Lower discontinuation rate • Increase enrolments 	<ul style="list-style-type: none"> •
Reviewers ("investigators")	<ul style="list-style-type: none"> • Membership: two internal (same university and School but external to the program under review (linguistics department): Dr. + MA) • Role: take a snapshot of the state of the program (strengths and weaknesses) and make recommendations ("opportunities for improvement") 	<ul style="list-style-type: none"> • Expert assessor
Evaluation questions	<ol style="list-style-type: none"> 1) What is the rate of discontinuation of the program? 2) When do students discontinue? 3) Do some languages have higher discontinuation rates? 4) Why do students discontinue? 5) How is [program name] viewed by students and staff? 6) Why do students enrol in the [program name]? 7) What issues arise during the [program name]? 	<ul style="list-style-type: none"> •
Review Model	<ul style="list-style-type: none"> • In-depth qualitative oriented study 	<ul style="list-style-type: none"> • Multi-method approach using qualitative and quantitative data • External approach

Methodology	<ul style="list-style-type: none"> • Data sources: students and staff perspectives (50/13); quantitative data from a previous review report and from student administration data base • Data collection methods: project phases (exploratory: identification of major issues preliminary report; aggregation and writing of final report); compilation of quantitative data; interviews (participants: selection criterion, rationale, distribution of number of students selected; procedure: information sought, interview style, focus of interview questions, pre-interview process, interview modes, location, length) • Limitations of methods, of review and tackling potential sources of bias commented 	<ul style="list-style-type: none"> •
Validity/Triangulation	<ul style="list-style-type: none"> • Quantitative (statistics from university database + previous review report) and qualitative data (variety of experiences of students, representativeness of participants across a number of variables # staff representing various languages and with diverse academic or managerial roles. However, uncertain about participation of casual teachers) • Collection methods: through interviews and compilation of quantitative data • Processes explained in detail 	
Stakeholders/audiences (for review)	<ul style="list-style-type: none"> • Students • Staff/Academics • Administrators 	<ul style="list-style-type: none"> • Primary (in this order): -Languages staff (casual teachers?) -Commissioner (+ management)
Reporting		<ul style="list-style-type: none"> • Written report (no evidence of any other reporting activities)
Utilization		<ul style="list-style-type: none"> • Doubtfully utilized due to anecdotal information indicating a negative reception of review recommendations

CASE STUDY 2 (Go8): Evaluation Fundamental Issues		
	Explicit	Implicit
Trigger(s)	<ul style="list-style-type: none"> • First 5-year review cycle • Alignment with institutional priorities 	
Evaluation Criteria ("context")	<ul style="list-style-type: none"> • University, Faculty (4 yr. specific period) and School (3 yr. specific period) Strategic Plans 	<ul style="list-style-type: none"> • Explicit only in name. Priorities in these plans are inferred through thematic analysis
Commissioner(s)		<ul style="list-style-type: none"> • Dean of the Faculty (constituted the RP) • University authorities
Evaluand	<ul style="list-style-type: none"> • School (integrated: two in one) 	
Purpose(s)	<ul style="list-style-type: none"> • Review overall performance of School • Learn about challenges and opportunities/ what elements 	<ul style="list-style-type: none"> • Alignment with institutional priorities

	“individuals felt should be retained or removed, changed or supported”	
Reviewers	<ul style="list-style-type: none"> • “Independent panel”: 2 internal (chair: “internal though independent” not affiliated to languages; + director of a “centre” associated with Asia) and 2 external (professors affiliated to languages (membership: name of univ. only) • Role: <ul style="list-style-type: none"> • Based on Strategic Plans, • “Engage in wide-range of discussions to learn about overall performance of School and elements which people feel should be kept/removed or changed or supported. • Report “where they wish” opportunities and challenges • Identify commendations (areas of strength) and recommendations (areas to be improved) • “evaluate”, “review”, “assess” 9 areas of the School 	<ul style="list-style-type: none"> • Additional stakeholder group, informants, providers of information redressed by an additional an unidentified “metareviewer”?
Evaluation questions	<ul style="list-style-type: none"> • Areas of study (policy driven?): <ol style="list-style-type: none"> 1) Governance structure (support strategic initiatives) 2) Internal organization (amalgamations) 3) Professional support 4) Financial performance and management 5) Management of curricula and teaching 6) Quality of research support and research performance 7) Quality of school’s support for internship programs (sustainability) 8) Strategic efficacy and sustainability of relationship between the School’s departments and programs and external stakeholders 9) Effectiveness of the School’s internal and external communication and marketing practices 	
Review Model		<ul style="list-style-type: none"> • Internal (university/School) (with no checks and balances): compliance with institutional policy and external (JIJO)
Methodology (Evaluation activities)	<ul style="list-style-type: none"> • Data sources: School’s “self-review report”; views from students, but mainly from staff and managers (33/8 and no casual teachers); additional documentation provided • Collection methods: meetings (interview schedule x 3 days), submissions (low response: 3); walking 	

	tour through facilities; opening for meeting requests (not commented)	
Validity/triangulation	<ul style="list-style-type: none"> • Priority in data sources: staff's perspectives (academics and managers excluding casual teachers) vs. students' perspectives. Acknowledged limitation with no explanation • Collection methods: diverse but absent/incomplete details as to procedure/analytical strategy 	
Stakeholders/Audiences (for review)	<ul style="list-style-type: none"> • Staff (academics and managers) • Students (undergrad./postgrad.) • Commissioners 	<ul style="list-style-type: none"> • Main audience for review and particularly review report: commissioners
Reporting	<ul style="list-style-type: none"> • Written Report • No evidence of other reporting activities 	
Utilization		

CASE STUDY 3 (Go8): Evaluation Fundamental Issues		
	Explicit	Implicit
Trigger(s)		<ul style="list-style-type: none"> • Compliance with institutional quality assurance policy requiring the conduct of a self-review for submission to the primary review (CS2)
Evaluation Criteria		<ul style="list-style-type: none"> • University and Faculty's strategic priorities: building a strong research identity, ensuring enrolments remain strong; participating in Faculty initiatives regarding infrastructure, curriculum management, and academic workload • The UEM (University Economic Model; consistent with conventional commercial concepts and management reporting practices): financial management tool; sustainable business management; education and research = core business + income generating
Commissioner(s)		<ul style="list-style-type: none"> • Same as CS2 (unknown identity ("dean"))
Evaluated	<ul style="list-style-type: none"> • School ("the School as a whole and not of individual departments.") 	
Purpose(s)		<ul style="list-style-type: none"> • Produce a report for the Review Panel to consider as a data source in the primary review • To make changes in the internal structure of the School ("recommends. from internal and external reviews have also affected the internal organization of the School")

		<ul style="list-style-type: none"> • Demonstrating fulfilment of Schools planned strategic directions in alignment with University Strategic goals “to improve its performance in areas of need and to establish collective objectives in medium term”
Reviewers		<ul style="list-style-type: none"> • Anonymous self-reviewer(s) • Role: prove alignment of School with University priorities through documents in Appendices and descriptions and evidence in written report <ul style="list-style-type: none"> - data gatherer, compile existing documentation (assembled by the Administrative Team (Organizational chart); university-generated sources? (e.g. description of roles/programs); departmental staff (departmental review reports) - interpret data - summarize findings pertinent to the six areas of School under review
Evaluation questions	<ul style="list-style-type: none"> • Areas of study (great resemblance to those in CS2, thus policy driven?): <ol style="list-style-type: none"> 1) In Introduction: Programmatic Development, SLC Strategic Objectives 2012-2015, Staffing And Student Load 2) Governance structure: School Board, Executive Committee, Departmental Structure, Internal Organization: Departments and Programs, Major Developments, School Coordinators, Effectiveness of the Interdisciplinary Programs, Challenges Programs and Departments Face 3) Professional support: Administration, Finance, Diploma of Languages, Language Study Centre, Infrastructure, Support Challenges 4) Financial management and performance: The University Economic Model, Impact on the School, Departmental Sustainability 5) Management of curriculum and teaching: Departmental Curricula Review, First Year Experience and Student Satisfaction, Language Pedagogy, Teaching Day, e-Learning Incubator, Contributions to PG Coursework Programs, Joint Teaching of Programs Across the 	<ul style="list-style-type: none"> • To what degree are these areas planned and operationalized so that developments are aligned with institutional priorities?

	<p>Faculty and other Tertiary Institutions, Exchange and other In-Country Study, BA Languages, Innovation in Teaching and Learning Activities, Diploma of Languages (Accelerated Mode), In-Country Studies at Peking and Fudan Universities</p> <p>6) Research support and performance: Research Support Officer, Research Advisory Committee (SLC Article Reading Workshop), Postgraduate Research Support Scheme, Conference Travel Support Schemes (Standard Conference Travel Support Scheme (since 2011); Top Standard Conference Travel Scheme (since 2013)), Research Outputs: HERDC and Non-Traditional Research Outputs, Research Income and Research Applications</p> <p>7) Relationship with external organizations and stakeholders</p>	
Review Model	<ul style="list-style-type: none"> • Self-review 	<ul style="list-style-type: none"> • Non-participatory self-review: an externally demanded task and a compliance exercise which requires to focus the study on particular areas, according to a quality assurance policy, but which gives room as to how the self-review process is conducted, which interestingly seems to adopt the same external character as the primary review
Methodology		<ul style="list-style-type: none"> • Data sources are mainly a variety of documentation (teaching review reports; research review reports, financial performance documents, summary of a School board meeting, etc.) • Data collection methods: compilation of documentation prepared by various stakeholders
Validity/Triangulation	<ul style="list-style-type: none"> • Variety of data sources?: Documentation not produced as a result of the review but existing documentation (from a past departments and programs' "extensive review" self-evaluation report and subsequent Strategic Statement and Action plan; previous year first yr. experience questionnaires results, financial information) • Data collection methods limited to compilation of documentation • Variety of perspectives: the writer's, a Pro-Dean Teaching and Learning, 	
Stakeholders/Audiences		<ul style="list-style-type: none"> • For review: The Review Panel of the primary review and the commissioners

		<ul style="list-style-type: none"> • For School: administrative team, students, staff including casual teachers, external stakeholders,
Reporting		<ul style="list-style-type: none"> • No evidence of other reporting activities other than the self-review report
Utilization		<ul style="list-style-type: none"> • No evidence

CASE STUDY 4 (Go8): Evaluation Fundamental Issues		
	Explicit	Implicit
Trigger(s)	<ul style="list-style-type: none"> • Five-year institutional review cycle 	
Evaluation Criteria (“context”)	<ul style="list-style-type: none"> • University [4 year period] Strategic Plan to build “a great research university” • Interviewee C: priorities in connection with building a reputation within the country and staying competitive 	
Commissioner(s)	<ul style="list-style-type: none"> • “offered by” (organized) Faculty of Humanities and SS administrators + • University administrators (PVC “finalised the terms of reference”; “appointed” reviewers) 	
Evaluand	<ul style="list-style-type: none"> • A language program (Diploma of Languages) and “language related courses” 	
Purpose(s)	<ul style="list-style-type: none"> • To identify strengths and weaknesses to help the program succeed • Interviewee C: To see how the program is doing in comparison to other programs around the country for the university to “position itself as a research” institution 	<ul style="list-style-type: none"> • To align evaluand with University priorities • Compliance with regulation • To capture talents that can help commissioners further to meet priority goals and for reviewers to make progress in their careers
Reviewers	<ul style="list-style-type: none"> • A panel of three members: Two external (one with language background; one “eclectic”: Law, insolvency, teaching reading legal materials in Asian lang.) + one internal (no language background) • Role: <ul style="list-style-type: none"> - “to serve on the review panel” - Comply with ToR: audiences to collect perspectives from; evaluand’s areas to review; criteria to review with; what not to recommend 	<ul style="list-style-type: none"> • Role: <ul style="list-style-type: none"> - Expert assessors: Gather data, form an opinion and provide recommendations - Informants: for commissioners to learn how other institutions go about their business of accomplishing priority goals to establish the university as a research institution - Potential assets - “we were not there to give the vice chancellor advice on what his statements and policies should be”.
Evaluation Questions	<ul style="list-style-type: none"> • Areas of study (Results chap.; various versions): 1. The Organization and Teaching of Language Courses 	

	<p>2. The Strategic Positioning of Language Study within the University</p> <p>3. The Diploma in Languages</p> <p>4. The Bachelor of Languages</p> <p>5. Promotion and Marketing of Language Study</p> <p>6. Study Abroad and Exchange Programs</p> <p>7. Language Pathways for Secondary Students</p> <p>8. Program Delivery and Information Technology Support</p> <p>9. Language Proficiency and Benchmarking</p> <p>10. Language Study across Faculties</p> <p>11. Language Teaching on a Cross-Institutional Basis</p> <p>12. Staffing and Management</p>	
<p>Review Model</p>		<ul style="list-style-type: none"> • Uncontrolled top-down internally initiated and managed, with audiences, including external reviewers, as providers of information/informants • Internally oriented and managed review (commissioners entitlement to set the rules) with external reviewers with role of informants whose findings become information to be used at convenience by commissioners
<p>Methodology</p>	<ul style="list-style-type: none"> • ToR: perspectives from a variety of audiences (“current staff, students and external stakeholders, including employers, alumni, and representatives of relevant professions and other relevant academic areas of the University”) • Review Process: <ul style="list-style-type: none"> - (3) data sources: (submissions + interviews for perspectives; documentation for background information about program) - data collection: for submissions via individual invitations (senior managers, student and staff representatives, external stakeholders), evidence of list of invitations in Appendix A and submissions received in Appendix B; for documentation (“Quality and Reviews” office), evidence of list of documents assembled in Appendix C; for interviews, evidence of schedule in Appendix B 	<ul style="list-style-type: none"> •

Validity/Triangulation	<ul style="list-style-type: none"> • Variety of data sources • Variety of data collection methods (deficient process/lacking information about process) • Imbalance in number of languages represented in student participants (1 German vs. 6 Japanese; no French but most academics French) • Students underrepresented (6/21) • No casual teachers • No explanation 	
Stakeholders/Audiences (for review)	<ul style="list-style-type: none"> • Primary audience: Commissioners • Program staff (no casual staff) • Students (underrepresented) • External stakeholders • Reviewers 	
Reporting	<ul style="list-style-type: none"> • Written report 	<ul style="list-style-type: none"> • Double-report writing by anonymous third party
Utilization	<ul style="list-style-type: none"> • Updated Implementation Plan • Interviewee C: Leadership: Level E position 	

Participants: Numbers and Diversity		Languages
Administrators	4	
Academics with administrative roles	4	Chinese (1) Japanese (1) French (2)
Academics	4	Spanish (1) German (1) French (2)
External Stakeholders	3	HS Principals (2) State Agency (1)
Students	6	German (1) Japanese (5)
	21	

CASE STUDY 5 (G08): Evaluation Fundamental Issues		
*Internal = from inside the university		
	Explicit	Implicit
Trigger(s)	<ul style="list-style-type: none"> • Ongoing program of quality assurance in Education • (I) Periodic reviews, embedded in the rules • Targeted review: the commissioning of the review was taken advantage of in order to make a personnel change, due to a “battle”, a “clash” between a dean and the HoS, to in turn push changes further • (M): A concern: “the School and the Faculty had a reason to be worried about the languages” 	
“Context” (Evaluation Criteria?)	<ul style="list-style-type: none"> • Within its context in <ol style="list-style-type: none"> 1. the College of Arts and Social Sciences 2. its linkages to other areas of the university (particularly the 	

	<p>languages and linguistics programs of the College of Asian and the Pacific)</p> <p>3. within the wider framework of the disciplines of Languages and Linguistics in the Australian university sector</p>	
Review Focus	<ul style="list-style-type: none"> • Academic Review 	
Commissioner(s)	<ul style="list-style-type: none"> • and conductors of reviews: The College Dean and Director of Arts and Sciences 	
Evaluated	<ul style="list-style-type: none"> • A School (of Language Studies) 	
Purpose(s)	<ul style="list-style-type: none"> • “to set future directions both in Education and Research” (to Schools and academic programs within the College) • to raise the leadership status of program in the nation • “to raise the education and research profile of the school with a view to taking its rightful place” 	
Reviewers	<ul style="list-style-type: none"> • 5 or 6 members (3 internal (1 = panel chair) and 2 external + Secretariat (internal)) • only one clearly indicated academic with language background (= M) - Management positions exceed academics = administrative orientation - Internal (counting the Secretariat) exceed external (4/2) - Affiliation to languages: 1 internal/1 external • Role: <ul style="list-style-type: none"> - “Consult widely” (“as appropriate”) - To examine areas in the ToR and provide suggestions for “measures or changes to practice in regard to the School’s operations” “where appropriate” - Report to College Dean and Director of College of Arts and Social Sciences. - Self-restriction: recommds. not only about “staffing issues” 	
Areas to be reviewed (Evaluation Questions?)	<p>-According to ToR:</p> <ol style="list-style-type: none"> 1. Organizational and governance structures (3 Q.) 2. Disciplinary base (3 Q.) 3. Range and delivery of undergraduate courses (4 Q.) 	

	<p>4. Capacity for further development of taught Masters programs (3 Q.)</p> <p>5. Higher degree by research (5 Q.)</p> <p>6. National and international research performance (5 Q.)</p> <p>7. Outreach to the community (local and nationally) (3 Q.)</p>	
Review Model	<ul style="list-style-type: none"> • No evidence 	<ul style="list-style-type: none"> • Compliance with a strong internal bureaucratic orientation whereby internal and external stakeholders, including review panel, participate as informants
Methodology	<ul style="list-style-type: none"> • Brief description in ES on data sources and collection methods: <ul style="list-style-type: none"> -<u>Submissions</u>: invitation to make a comment (inconsistency) + submissions received (list of 14 staff/students/group-community who sent in submissions) -Embassies (number and names) / wide range of cultural organizations (no names/number) -From three groups: School, academics (incoherent?) and students • <u>Meetings</u>: (process?) with staff from the School and the university, external stakeholders, students • <u>Information Session</u> w/ students "held by a member of the panel" • Tour of facilities • Analytical strategy? 	<ul style="list-style-type: none"> • Confusion due to censored information in "Submissions received" (from [...]): Do submissions include Academic review reports? Censored or not indicated in this section but the report states that the School did make a submission
Validity/Triangulation	<ul style="list-style-type: none"> • Diversity: data sources/collection methods (= diversity of perspectives) <ul style="list-style-type: none"> -Submissions received (no description/inconsistency): -Internal*: 7 (School), 4 (College of Asia and the Pacific), 1 (multiple: by linguists) • Balanced academic/managerial roles? Unable to assess; information covered -Students: only number (= 5) -External: 1 individual ("on behalf of" verbal submission then written after the review; names covered); 5 embassies (names + information covered) -Meetings: (scant description; inconsistent inform.) -“Staff attending the review” (description: format (group/individual-participation of one of the individuals “twice”),): names of 4 individuals (1 (School x2), 1 (School = another School), 1 (Asia and the Pacific), 1 (Research and Graduate Studies); participants in groups: “the staff of the School” -Students: number of students who met with the panel (=12); inform. session w/ students by blacked out name (purpose of session: inform about review and receive students comments about experiences at School) 	

	-Community: someone's name who "attended <i>in</i> the Community session"	
Stakeholders/Audiences for review	<ul style="list-style-type: none"> • Commissioners/organizers • Informants (review panel including <i>secretariat</i>) • University staff • School staff • Students • External stakeholders (embassies) 	<ul style="list-style-type: none"> • Main audience for review: commissioners ("the Review Panel will report to the of the College Dean [...])
Reporting	<ul style="list-style-type: none"> • No evidence except written report. • School staff representativeness informed during review (discussions with reviewers) about changes that were sought 	<ul style="list-style-type: none"> • Double-report writing by anonymous writer
Utilization	<ul style="list-style-type: none"> • Evidence in interview data, not in report 	

CASE STUDY 6 (Non-Go8): Evaluation Fundamental Issues		
	Explicit	Implicit
Trigger(s)	<ul style="list-style-type: none"> • Context leading up to review: (in Self-review report): Announcement of program closure (= outcome of a meeting) "later replaced by a commitment to a process of consultation and review" • Commissioner's Review notice: 4 triggers (1) low enrolments and high attrition rates (2) financial costs (3) lack of implementation of recommendations from past reviews (4) "community concern about earlier proposed closures" • Self-review report's outcome: false triggers and assumptions 	<ul style="list-style-type: none"> • True triggers: -the "campaigning", the "huge resistance, "the "media coverage" and the "noise" which caused "embarrassment" (A) [...] So decision first, and met the huge resistance including the media coverage. That made him to do review [...]
Evaluation Focus	<ul style="list-style-type: none"> • Primary review + ToR for self-review: -review perceived as "firmly framed around finances" -(K): review used as pretext not only to deal with finances but specifically to push false argument, that is, a deficit in finances: K64. [...] And one of the issues, of course, was the financial aspect of it. The driver for the review was, you know— it was supposedly costing money, losing money. 	

	<p>Turned out, of course, it wasn't losing money</p> <ul style="list-style-type: none"> • Self-review: ToR's focus + proving misinformation and incorrect assumptions in Review Notice, program's achievement of strategic goals and internationalization of education policy, program's alignment with Federal Government policies; extensive connections with outside world; vision for future development, contribution to campus culture 	
Evaluation Criteria	<ul style="list-style-type: none"> • Viability 	<ul style="list-style-type: none"> • None: predetermined decision
Commissioner(s)	<ul style="list-style-type: none"> • dean of Arts, and responsible for the eventual closing down of whole language program at university <p>A208. This was er the dean] who actually shut down the program so that dean's name is (P22), he is doing the faculty so—</p>	<ul style="list-style-type: none"> • dean not only perceives himself as member of top univ. authorities but speaks on behalf of Univ. <p>"we have all the information we need regarding other aspects [...]"</p>
Evaluand	<ul style="list-style-type: none"> • An undergraduate language major 	
Purpose(s)	<ul style="list-style-type: none"> • Self-review's 4 purposes to comply with ToR but partially due to biased ToR: <p>(1) answering questions "presented [...] to staff of the [language] area"; "address each of these questions [...] with extra data and information in Appendices"</p> <p>(2) addressing "some contentious and erroneous statements in the "Triggers" statement in the review notice"</p> <p>(3) discuss wider context (Government policy) "where it affects matters of future viability"</p> <p>(4) address specifically ToR 2, that is, to make recommendations about future of program + suggesting options for students</p>	
Reviewers	<ul style="list-style-type: none"> • 2 groups of reviewers: <p>1. Review Panel: 3 members; no affiliation to languages (1 external –Emeritus Professor-, chair of panel; 2 internal –Dean, Faculty of Business + Associate Dean, Faculty of Education)</p> <p>-(C): negative perception about the composition of RP; disapproval of lack of language representation</p>	<ul style="list-style-type: none"> • Role (self-reviewers): <p>1. informants: gather information for the Review Panel to make a judgment and pass their recommendation</p> <p>a. for the commissioners to make a decision on whether the program is or isn't financially viable.</p> <p>b. to accommodate the commissioners predetermined</p>

	<p>2. Self-reviewers: 1: Head of department (Dr.), 1 member of departmental staff (Dr.)</p> <ul style="list-style-type: none"> • Role: -2 tasks: <ul style="list-style-type: none"> 1. study viability (enrolment numbers + financial outcomes) 2. provide recommendations about future of program and options for students -compliance with expectations in the ToR and beyond “We have responded in detail to the questions and allegations in the review notice, documented current plans, and developed new plans.” 	<p>decision to close the program (according to verbal data)</p> <p>2. advocators of the program to prove the inaccuracy and misinformation on which the commissioning of the review was grounded and eventually save the program</p> <ul style="list-style-type: none"> -to protect their dignity as professional language educators • Role (Review Panel) -(see above +) consider, study contents of self-review report, provide data to self-reviewers
Evaluation Questions	<ul style="list-style-type: none"> • Three questions in ToR (same as primary review?): <ul style="list-style-type: none"> (1) actions taken to implement recommendations from past reviews, particularly in connection with high attrition, small enrolments, the problem of class-spread over a number of days and teaching costs significantly exceeding income from student fees (2) plans to increase enrolments, to improve retention and reduce costs (3) plans to eliminate current deficit 	<ul style="list-style-type: none"> •
Review Model	<ul style="list-style-type: none"> • Self-review 	<ul style="list-style-type: none"> • Primary review: external/internal evaluation; highly external quality evidenced in how self-reviewers perceive the self-review as a “submission”
Methodology	<ul style="list-style-type: none"> • Data sources + data collection methods: • Self-review: <ul style="list-style-type: none"> -an “extensive review” -“income sources”, 3 past review reports, “modelling costs and benchmarking”, letters (48 pgs.) from external stakeholders and alumni -Ongoing investigations: “true retention rate for the major”; first yr. student numbers over 8 yrs.”; surveys on motivation for studying Japanese -Benchmarks: bet’n Japanese courses attrition figures @ university and other “education language education contexts” 	

	<p>-Distribution of self-review tasks: A266. Sure]. Yeah. This one, I made a 2 version. Everything I wrote is my work and anything, other than financial information— Q267. You were the convener of the program at that time.</p> <p>A268. Yeah, I wrote the whole thing so um there are a few things like er— [...]</p> <ul style="list-style-type: none"> • Methodological limitations: obstruction to data (from interview data) • Primary review: documentation (self-review report, figures, cost calculation, student numbers) 	
Validity/Triangulation	<ul style="list-style-type: none"> • Self-review: diversity of perspectives (alumni, external stakeholders, of data sources and data collection methods) • Primary review: only documentation, evidence of sharing of documentation with program staff (self-reviewers), no evidence of meetings with program staff, no evidence of students involved; serious issues threatening fairness of process 	
Stakeholders for Review	<ul style="list-style-type: none"> • Self-review: program staff (self-reviewers), external stakeholders, alumni, Review Panel, commissioner • Primary review: commissioner 	
Reporting	<ul style="list-style-type: none"> • Self-review: 1 report • Primary review: no evidence 	
Utilization	<ul style="list-style-type: none"> • Primary review: decision taken and implemented: language major closed, all languages closed 	

CASE STUDY 7 (Non-Go8): Evaluation Fundamental Issues		
	Explicit	Implicit
Trigger(s)	<ul style="list-style-type: none"> • Standard formal internal periodic procedure: "Program Evaluation and Re-Accreditation Review Report" 	
Evaluation Focus	<ul style="list-style-type: none"> • emphasis on marketing, promotion, enrolments, attracting students 	
Evaluation Criteria	<ul style="list-style-type: none"> • University Strategic Plans: 	<ul style="list-style-type: none"> • Attracting students
Commissioner(s)	<ul style="list-style-type: none"> • Three points of internal accountability: 	<ul style="list-style-type: none"> • University (no names)

	(1) Head of School (2) Dean (Teaching and Learning) (3) Division PVC	
Evaluand	• Bachelor of Arts (Languages and Intercultural Communication)	
Purpose(s)		<ul style="list-style-type: none"> • to evaluate and Re-accredit program internally • establishing program's directions based on predetermined University Strategic plans • to demonstrate program's compliance with Strategic priorities, particularly demonstrate commissioners efforts to ensure student numbers are sustained and increased, that is, that program is viable
Reviewer(s)	• Teaching team (no names)	
Evaluation questions	<ul style="list-style-type: none"> • Areas reviewed: 1) Program viability 2) Retention and success 3) Current student satisfaction 4) Graduate satisfaction and outcomes 5) External stakeholder satisfaction 6) Academic design 	<ul style="list-style-type: none"> • Is the program aligned with broader priorities? • What activities is the program engaged in in order to align program with priorities, particularly with ensuring student numbers are maintained and increased?
Review Model	<ul style="list-style-type: none"> • Participatory Self-review: <ul style="list-style-type: none"> -data sources and data collection methods undertaken by the teaching team -14/34 "we" references: on data analysis thus suggesting collaboratively gathered and studied -staff's engagement in the action plans to implement after the review, as recommendations 2, 3, 4, 5, 6, 11, and 12 in the "Summary actions/changes proposed" chapter prove 	<ul style="list-style-type: none"> •
Methodology	<ul style="list-style-type: none"> • Data sources and data collection methods: - "program data, and data obtained from collection methods such as benchmarking and focus group discussions". -feedback from students, "both ongoing and in focus groups", as well as feedback from "external stakeholders". -Further anecdotal data sources suggested: "While anecdotal evidence suggests [...]" under "Graduate destinations and employment" 	

	<ul style="list-style-type: none"> • Analytical strategies/focus/use <p>1. to compare program with similar programs -to learn about likely international markets -Market analysis for market position of program: analysis and comparisons with CEQs; analysis of competitors with similar programs</p> <p>2. To study program coherence: -Mapping analysis within and across all courses focused on: -teaching and learning arranges. -objectives -content -assessment</p> <ul style="list-style-type: none"> • Methodological limitations (reg. data sources): -reliability of data (no record of some TER scores -no data for explaining a finding about student profile, equity and inclusivity -problematic comparisons of scores across programs of other univs. for the way data is reported -data for Grad. Satisfaction provided by very few students 	
Validity/Triangulation	<ul style="list-style-type: none"> • Variety of data sources • Variety of data collection methods • Variety of stakeholders' perspectives 	
Stakeholders for Review	<ul style="list-style-type: none"> • Students • Teaching team • External stakeholders • Commissioners (three points of accountability) 	
Reporting	<ul style="list-style-type: none"> • No evidence of other reporting activities other than this review report 	
Utilization	<ul style="list-style-type: none"> • Implementation plan designed 	

Appendix 5: Case Studies Coding

CODING

CASE STUDY 1 (Go8) Year: 2005	
THEMES FROM REPORT	
Most themes found related to EFIs = focus of report on ensuring a methodologically valid process, thus professional/scientific approach	
1. Limitations of study	<ul style="list-style-type: none"> • Limitation of previous quantitative analysis: “One possible concern with quantitative analyses [...]” • Analytical limitations (impossibility to theorize about “individual languages” since “student numbers across languages were not equal” but general trends identified • Lack of evaluation criteria: Unable to know reasonable rate since no data from other degrees to compare it with “Since no data is available about discontinuation rates from other degrees and diplomas it is difficult to project what level of discontinuation rate would be reasonable, particularly because of the “add-on” nature of the [program name] might make it generally more susceptible to discontinuation.” • potential sources of bias in data collection: <ul style="list-style-type: none"> • “Participants self-selected to some extent by making themselves available, possibly because they held strong views about their language studies” which reviewers tackle by reaching students with “a wide variety of experiences”.
2. Reviewers’ judgments	<ul style="list-style-type: none"> • informed?/Research/references not always used/quoted <ul style="list-style-type: none"> -“Some discontinuation is unavoidable”; “We acknowledge that there will always be some discontinuation [...]” -“Another point sometimes raised with discontinuation [...]” -“Having sympathetic instructors and an encouraging non-threatening class environment also assists [...]” -“In line with previous research, learning languages also brought vocational and personal development and [...]” -“[...] is an issue that has emerged in many settings where communicative language teaching is prevalent [...]” -“[...] familiarizing staff with such fairly novel teaching approaches is one of the key recommendations below.” -“Nothing annoys language students more than [...]” - [...] but a discontinuation rate around [...] of a cohort is far too high.” -“There is no compelling reason why [...]” -“The timetable should reflect that [...]” • Coarse judgmental language? <ul style="list-style-type: none"> -“A placement system that assigns [...] is antediluvian [...] and irresponsible [...]” -“Class sizes of 20-30 students are pedagogically disastrous [...]”
3. Interpretation of data	<ul style="list-style-type: none"> • E.g. “By this students mean that [...]”
4. Contextualization of findings	<ul style="list-style-type: none"> • E.g. “Reasons for discontinuations should not be isolated but viewed in conjunction with [...]” -student cohort’s view (expendable, supplementary degree) -transition change from secondary to tertiary education
5. Future research	<ul style="list-style-type: none"> • E.g. -“Further research may reveal too that these students [...]” -“Whether students took advantage of [...] remains to be seen”
6. Theoretical discussion	<ul style="list-style-type: none"> • theoretical discussion on discontinuations with no bibliographical references: unknown reasons, external factors, “good discontinuation”, the system

<p>7. Methodological procedures</p>	<ul style="list-style-type: none"> • Brief description of data collection methods and data sources in ES and nuanced description in Methodology chapter: project phases; quantitative data; interviews: <ol style="list-style-type: none"> 1. participants: Students: selection criterion and rationale, representativeness of participants, distribution of numbers of students; 2. procedure: process of interviewing (interview modes, interview style, description of content and rationale of interview questions, place, length, ethical process in preparation); Staff: variety (further inferred from Results chapter: students don't differentiate in their comments among teachers by their "employment terms" may indicate "casual staff" accountable for the bulk of language teaching may have been part of Staff interviews but evidence only shows staff interviewees represent various languages; two heads of department; head of the School. -Calculation method of discontinuation rates described
<p>8. Review focus</p>	<ul style="list-style-type: none"> • Discontinuations + strengths and weaknesses (student numbers + improvement) <ul style="list-style-type: none"> -Summary of main findings (positive: positive perception of program; negative: high discontinuation rate) -Purpose of recommendations: reduce discontinuations, increase enrolments, maximize students' time -One main recommendation: inform students before enrolling (to avoid discontinuations) -Purpose of review: understand reasons behind discontinuation rate (why students discontinue) + take overall picture of program -Lower discontinuation rate by keeping students (by improving some areas) and increasing students (by improving other areas: information to students/marketing) -Main implication of findings: lower discontinuation rate and increase enrolments -Reviewers' role: snapshot of program, strengths and opportunities for improvement -Description of program: number of points, range of languages, length, general aims but emphasis on <ul style="list-style-type: none"> • enrolments, specifically enrolled "fee-paying students" versus "sponsored places" both from overseas and Australian • popularity of languages in terms of "enrolments" during a four-year period -Evaluation questions rate of discontinuation of program; Why students discontinue; students and staff's perception about program; issues in program -High number of recommendations (24): focused on improvement in three categories (4, 10 and 10): thoughtfully balanced distribution with the "urgent" ones least in number for immediate and prompt effect • Focus on primary program's stakeholders <ul style="list-style-type: none"> -Focus on participants' perspectives: interest in improving quality of educational experience): 63 participants interviewed: students (50) outnumber considerably staff (13) -Students and staff's perspectives in detail: positive perceptions + concerns, needs, what students would like, want, etc. and suggestions for changes or improvements in Results chapter; reasons why students discontinue: general trends.
<p>9. Program's accountability</p>	<ul style="list-style-type: none"> • Recommendations: for academics and administrators: distributing accountability/responsibility among different stakeholder groups -Administrative (10 summarized in 7 due to 3 on "flexibility"(R3) = facilitating access): <ol style="list-style-type: none"> 1. smaller class sizes (20-30 disastrous/ around 15 ideal for conversation and interaction)

	<ol style="list-style-type: none"> 2. introduce Certificate of Modern Languages (encourage the 84% who leave to stay and get qualification) 3. flexibility (a) degree: stand alone, widen participation; b) timetables flexible: spread across week, no consensus for night classes; c) more flexibility: 2nd and 3rd year: mix and match streams) 4. more higher level subjects (separate post-VCE more from less advanced) 5. information sessions (induction/follow up to students who are currently under-informed on administrative details) 6. suitability of teaching spaces (adequate conditions for teaching languages) 7. promote overseas opportunities (increase opportunities); team teaching (two tutors per class: exposure to accents and increase interaction) <p>-Academic:</p> <ol style="list-style-type: none"> 1. conversation and communicative skills (vs, translation and rote learning; small class sizes indispensable) 2. appropriateness of teaching styles (associated with methodologies from target country considered authoritarian, old-fashioned) 3. elaboration of measurable outcomes (at each entry level; objective yardstick to measure program success) 4. subjects integrated (no current unity between subjects) 5. slow down pace (fast-tracking beginners leads to loathing and discontinuation) 6. frequent small assessments (vs. rare and large ones); 7. more contemporary materials (to develop practical language abilities) 8. language learning outside the classroom (e.g. partner programs run by the appointed program coordinator) 9. webcraft/Blackboard (class attendance vital but back up of material useful) 10. teaching media (vs. computer aided instruction: grammar computer programs; more use of language labs)
<p>10. Focus on student numbers</p>	<p>•Although purpose of quantitative study not revealed, “follow-up” nature of current review suggests one of the areas studied, if not main driver of quantitative study may have been focus of present review = “student discontinuations”. Therefore, possible to claim first review may either have been conceived too to lower discontinuation rate or that high discontinuation was an outcome ensuing from quantitative study</p>
<p>11. Trigger: staff’s ongoing refusal to address discontinuation</p>	<p>•Main trigger: evidence below reveals commissioning of review with corresponding modifications may have been triggered by need to enhance evaluation utilization, that is, an opportunity to gather further evidence to substantiate the findings from previous quantitative study in an attempt to increase staff’s commitment to changes that were perceived as needed by commissioners</p> <p>•Change in review approach: From quantitative to qualitative review</p> <p>-why conduct a qualitative study when discontinuation rate may have ensued from quantitative study?</p> <p>-- dates from both reports showing three-year lapse between the two: indication that no changes were implemented after quantitative report and that the issue of discontinuation persisted.</p> <p>--negative reception of findings from quantitative report may have prevented commitment to actions. Only further research could shed</p>

	<p>light on opposition, but commissioning of consecutive review suggests commissioners may have perceived quantitative orientation of first review potential cause of negative reaction, thus commissioning a qualitative one.</p> <p>-- two further evaluation issues changed:</p> <ol style="list-style-type: none"> 1. Reviewers: "preparer" of the first report = individual not affiliated to languages¹²⁰ to reviewers in present review with AL background, 2. Source of commission from "the Faculty's Academic Programs Office" in the first review to a closer-to-home commissioner = Head of School <p>--Available anecdotal data suggests new methodological approach not issues affecting evaluation utilization since findings from present review were not well taken either ("they hated it")¹²⁰.</p> <p>• Reasons for staff's lack of commitment to addressing discontinuation after the quantitative study:</p> <p>1. Reviewers, non-experts: legitimacy issue?</p> <p>- "Investigators": identity, tasks, academic roles. Do the reviewers' academic positions, particularly, that of the MA student, and their admittance of no expertise in evaluation ("investigators" vs. reviewers/evaluators) may jeopardize legitimacy of review? May these contribute to perception of review as non-legitimate?</p> <p>2. "discontinuations" perceived as a concern by commissioners but not by program staff, and therefore, second review understood as an instrument to force staff to deal with an issue that program members were not quite ready to address: enrolments associated with finances perceived as non-legitimate in an educational institution</p> <p>But, persistence in non-commitment after qualitative review, why?¹²⁰</p> <p>Commissioner's interest used as pretext for refusal to commit to utilization</p>
<p>12. Urgency for improvement (discontinuation rate + other)</p>	<ul style="list-style-type: none"> • First milestone: cut rate by half in near future by keeping or/and increasing enrolments • Commissioning: following up a quantitative in nature report with data on discontinuation rates • Summary at end with set of urgent recommendations and other improvements (administrative and academic) <p>-Urgent recommendations (improvement/promotion):</p> <ol style="list-style-type: none"> 1. appoint a coordinator (advertising: to increase enrolments; advising and supporting staff and students; part time: + other responsibilities: language class coordination for HLC (Horwood Language Centre) 2. website (run by program coordinator) 3. new placement (abandon VCE/non-VCE dichotomy: criticized by reviewers: antediluvian/irresponsible; specific suggestions for slight, major and cutting edge improvements); 4. professionalization of language teaching (offer professional development program: short-term/long-term approach; current graduate certificate in teaching inadequate)

<p>CASE STUDY 2 (Go8) Year: 2014</p>	
<p style="text-align: center;">THEMES FROM REPORT</p>	
<p>Values</p>	<ul style="list-style-type: none"> -Collegiality -Positive work culture, good working relationships - Value of teaching and research (supported by senior management in Faculty: Stating the obvious?)

Alignment with Strategic priorities (associated with financial viability, sustainability: cost-effective / income generating)

- **Strategy driven by efficiency:** One school manager for two schools: doesn't make it to a recommendation. Reviewers avoid recommending what implies the increase of resources or shyly and/or euphemistically request it. In this case, this inconvenient situation is described as "it leads to a divided focus".

-**Strategic program development** associated with **control** and ensuring **resources are sought:** e.g. creation of a curriculum group: for oversight and ensure people have a deep understanding of changing environment (= resources may not always be there to support curriculum)

-**Academic matters such as curriculum through strategic lens:**

- Curriculum means management of curriculum, adapting programs to strategic goals with a focus on viability and sustainability. E.g. in recommendations: strategic plan to ensure curricula are sustainable

-**School's institutional reputation a challenge. Recognition gained through:**

- **marketing internal income-generating programs**

E.g. accelerated mode language program = needs to be known: to help shape broader institutional discussions = so other departments follow model

- **engaging in programs.** E.g. creating wider access + opportunities for students (= increasing enrolments)

- **Diversity:**

- **Success = Balancing diversity with sustainability:** Breath of languages, range of disciplines, diversity of programs, range of provision
- **Diversity of curriculum** approach assoc. with **innovation (making teaching more cost effective/attractive to maintain/increase enrolments)**
 - e.g. **accelerated mode program** for widening access (access to languages), a sign of innovation and strategic focus. This program associated with income with a risk of becoming a "cash-cow" business
 - e.g. **e-learning** (incubator: e-learning start up solution for more **cost efficient e-learning**): run administratively
 - e.g. promoting teaching innovations (**blended learning** (face-to-face classroom teaching combined with online teaching) + educational technology) = technological approach to teaching praised
 - Curriculum reform** assoc. with e-learning initiatives
 - Professional development** = educational innovation

-**Concerns: sustainability** issues of some programs = are they **financially robust?** To what extent have they proved to be **financially viable?**

-**Context** (= Evaluation criteria): Strategic Plans of the University, of the Faculty and of the School relative to a particular year period

-**Appendix 3 (non-confidential submission):**

- **a concern:** study area "under (financial) threat", in need of resources, particularly staffing. Letter argues the need to continue the program alluding non-financial arguments: importance of region, tradition, a priority in University's international plans.

	<ul style="list-style-type: none"> • Causes for lack of resources: was allowed to run down, failure of resource rationale for cross-subsidization <p>-Cost-effectiveness: Integration (two groups of languages into one School); amalgamations, relocation; program description's focus on structural changes</p> <p>-Strong level of external engagement praised: significant resource for the School "to support effective curricula and strong research agendas", that is to serve "the strategic direction and priorities of the School"</p> <p>-Praised efforts to strengthening research culture assoc. with finances: e.g. gains in research performance, above all "significant growth in grant income"</p> <p>-School's strategic planning to ensure student numbers do not decline</p> <p>-Cross-listing with relevant departments (to increase enrolments?)</p> <p>-Marketing: attracting and keeping students (enrolments)</p>
<p>Focus of Strategic Priorities (non-financially related? to fulfil other University priorities such as rankings, etc.)</p>	<p>-Priority: Research</p> <ul style="list-style-type: none"> • Praised efforts to strengthening research culture: many activities which contribute to "a growing recognition of the need to prioritize research outcomes" • Interdisciplinary networks, transnationalism network and language and identity research network • Strong disciplinary clusters, broad research strategies and priorities, administrative support to researchers • Quality of research • Expectations for academics (in staff's assessments) for research • Avoid academic isolated silos (= networking, collaboration desired) <p>-Other priorities:</p> <ul style="list-style-type: none"> • Reputation: internal and external; national and international • Gender balance praised • Protecting disciplinary identity • Ensure representation of particular countries • Communication: of students' views and communication processes at various levels • Understanding the value of language study across campus • Vision/position of School and contribution to university profile and the wider community • Reviews • Academic governance • Balance and coherence of content
<p>Defensiveness, fear</p>	<p>- Concern: viability</p> <ul style="list-style-type: none"> • Repeated praising (x3: ES, Findings, Summary/Conclusions) about School's success in maintaining finances and understanding and working towards managing to remain "financially strong" • after praising School's "strong suit of programs" anonymous writer states: "The panel would encourage the University to ensure that this diversity continues to be supported and maintained" • Timid language when requesting resources: "would ideally require an increase of one member of staff" • Repeated praising (leadership, innovation, for internal recognition) of accelerated mode program for languages

	<ul style="list-style-type: none"> • Hyperbolic language (overpraising) in relation to finance: Financial management and performance: obviously strong <p>-Concern: student numbers</p> <ul style="list-style-type: none"> • Threat to School: decline in student numbers due to fee deregulation (universities free to charge students what they want which puts pressure on students to jet in jet out of university to get jobs as quickly as possible to pay back student loans). Thus strategies to consider to attract students: overseas study, inform: skills students gain and career pathways.
Compliance/eagerness to compliance/praising compliance	<p>-Repeated praising of cooperative behaviour of participants in review</p> <p>-Positive response to change due to acceptance of responsibility for viability and of relationship between academic integrity and operational and strategic imperatives. E.g. discontinuation of Masters coursework</p> <p>-Financial management and performance: obviously strong thanks to income from strong teaching programs, research strengths, and cross-subsidization (public and community engagement). <i>Changing character of these patterns accepted</i></p> <p>-Leadership of the school praised in front of challenges (languages and integrative school model) (= successful in complying with cost-effectiveness strategies)</p> <p>-Collegial culture with a view to remain financial strong by carefully managing its programs</p> <p>-In commendations:</p> <ul style="list-style-type: none"> • Proactive approach to ensure sustainability • Understanding and responsibility to ensure viability of curriculum • Effective financial performance
Threatening when no compliance	-School's administration: stable in so far as it understands the need to adopt new approaches (= expectations in performance management)
Association of reviews with "viability and/or sustainability" of curricula	<p>-Financial strength = no reviews: A program "has been able to avoid serious reviews" due to their external funding</p> <p>-Sustainability at risk = review: Review of potentially unsustainable programs</p>

CASE STUDY 3 (Go8) Year: 2014	
THEMES FROM REPORT	
A broader theme ensuing from the themes below can also be TOP-DOWN system where academics can express but are not part of the decision-making process	
Association of review with financial performance	<p>-Since this review is of the whole school, it is worth noting that three departments have returned significant deficits results</p> <p>- Departmental curricula reviews: in numerous cases cutting low enrolment units/units with no strategic relevance for curriculum. Overall result: leaner/more streamlined curricula</p>
Alignment with institutional strategic priorities	<p>-Strategic objectives and purposes to be "aligned with both the University and Faculty strategic plans" regarding priority areas:</p> <ul style="list-style-type: none"> • research: profile, expectations • service: equity among academics, foster productive relationships • curriculum development: healthy and responsive curricula to changes (discipline and students' interests)

	<ul style="list-style-type: none"> • language teaching: balance language/non-language units; teaching interests • technology: for our purposes and students' needs <p>- Governance Structure's strategic function = ensure alignment with broader goals:</p> <ul style="list-style-type: none"> • School Board: development of Operational Plan • HoS: strategic planning in line with planning at Fac. and Un. levels • Executive committee: practices informed by wider institutional issues and perspectives • Depart. Chairs: strategic and operational duties • School coordinators: support strategic initiatives <p>- Curriculum content subject to strategic priorities: Strategic "major developments": Japanese and Chinese Buddhist specialists: = appoint first and redesign curriculum second; align first (recruitment in strategic fields) and adapt to curriculum second</p>
<p>Demonstrating strategic priorities tackled guided under efficiency principle (- Efficiency: budgetary discipline (short term gains, keeping costs down e.g. online courses), income generating (research (grants), innovation (keeping students satisfied to attract/retain/increase enrolments))</p>	<p>-Strategic Plan assoc. with UEM (University Economic Model):</p> <ul style="list-style-type: none"> -Univ. Strategic Plan 2011-2015 -Financial management tool -Encourage sustainable business management -Education and research = core business & income generating (income allocated in full to units that generate it) -Consistent with conventional commercial concepts and managing reporting practices: operating costs allocated (how much service and infrastructure they consume) and charged a levy (fund Univ. wide strategic initiatives) <p>-“contexts for understanding current organization of the School”:</p> <ul style="list-style-type: none"> • reconfiguration of the School (to secure administrative efficiencies) • merging of disciplines, increased number of Schools (3 to 5) in Faculty (from Faculty of Arts >> Arts and Social Sciences) <p>-programmatic development: additions, eliminations, splits, renaming, moving, becoming, joining, developing into; HoS's retirement, and appointment of new HoS</p> <p>-Interdisciplinary programs (effectiveness/success = degree of efficiency):</p> <ul style="list-style-type: none"> • Asian Studies: major success: number of units, annual enrolments, number of language departments, postgraduate units' enrolments • European Studies: redesigned –appointment for that task; Challenge: reach similar level of success as Asian Studies. Reasons for lack of success: culture units in target language; undergraduate units do not count towards major; competition with ICLS (Intern. and Compar. Literary Studies) <p>-Staffing/Student-staff ratios (SSR): reduction in staffing three departments with significant deficit results: comprehensive implementation plans in place for increasing revenue through strategic increases in student load; successful fundraising campaign (\$2m) of one of such departments</p> <p>-Diploma of Languages (accelerated mode: 4 langs.; 29 units (24 blended learning + 5 fully online;</p> <p>-Administrative support: administrative positions (small in number but unified and versatile; provide adequate coverage)</p> <p><u>-academic positions</u> (need support for Teaching and Curriculum)</p>

-**Language Study Centre** managed by the School (librarian made **redundant** after reconfigured: digitalization of audio recordings; videos made available via Blackboard; other material via Library)

-**Infrastructure:** office spaces; one room for e-Learning Incubator (Diploma of Languages (accelerated mode); rooms given to the South-East Asia Centre; pressed for space

-Support challenges: centralized administration: greater flexibility and coverage; feeling of belonging and support; low turnaround// university restructuring' challenges: communicating changed procedures and maintaining levels of support (e.g. tasks devolved to School level with no additional resources)

-**Departmental curricula reviews:** in numerous cases cutting low enrolment units/units with no strategic relevance for curriculum. Overall result: leaner/more streamlined curricula

-**First yr. experience and student satisfaction:** committed to improvement through participation in Faculty initiatives (e.g. increasing Blackboard sites and blended learning environment: e-Learning Incubator); commended by Pro-Dean Teaching and Learning: Blackboard sites and reporting on Unit of Study Evaluations: very positive and high response rate = students' high rate of class participation and engagement)

- **Language pedagogy:** very active committee: priority areas, improve learning needs and **e-Learning presence**, reflective and collaborative culture, support research, information sessions, workshops, conferences,
- **Teaching day:** presentations, workshops and discussions; improve lang. pedagogy, sharing best practices (e.g. past one: 4 guest **speakers:** one national (**computer-assisted language learning**); three international: 1 (ESL) + EAQUALS; 2 **Arabic** language)
- **e-Learning Incubator:** essential equipment, collaborative technology, pedagogical and technical support; **redesigning curricula for blended learning**, particularly the Diploma of Languages and **four innovative blended learning projects** in other langs.: energy and focus to **overall blended curriculum re-design** process; 4 functions of e-Learning Incubator: one-stop-shop space; interaction and collaboration; sponsor and support life cycle of blended learning projects; know-how and equipment audio/audio
- **Contribution of PG coursework programs: elimination of low-enrolment units and coursework programs** to free up resources for successful UG ones; retaining high enrolments ones (part of other majors)
- **Joint teaching programs:** long tradition across Faculty or other institutions: e.g. **successful unit with regular enrolments** higher than 200 and **Chinese Business** quickly gaining popularity and increasing enrolments; new agreement with another university (exciting chapter: students must comply, Commonwealth supported (subject to proving with document. vs. not supported (charged fees e.g. international students, etc.)
- **Exchange and In-country study:** encouraged; credit counts towards a degree; exchange institutions (over 260) best and most prestigious in the world; reputable; positive testimonies from students.

-**BA languages:** four years; ATAR cut off; 2 majors or 3; **one in-country study semester or two a must**; grant of \$5000

-**Innovation in teaching and learning activities:** commitment to constant improvement; leading the way: **accelerated program + in-country study at two Chinese universities**

-Diploma of Languages (accelerated mode: accelerated mode Diploma (new and efficient in modern univ.) (Spanish and 3 Asian); for what students (**widening participation:** students with units free for lang. and grad. and postgraduate; structure of program (intensive); duration (1 yr.); typical structure:

- 1 common unit (online/intensive; intercultural communication)
- 1 specific country unit (English; online; background knowledge; specific disciplines: health, business, engineering)
- 2 level-1 units (intensive during semester break)
- 4 upper-level units (semester/winter-summer school format; in country or combined)
- Language learning beginners (relevant thematic areas, e.g. practical lang. associated with corporate, public service and research engagement (telling time, phone calls, etc.)
- In-country units (with same curriculum at partner university)
- Internship units (cross-listed at upper level or substituted by unit in-country to increase relevance and flexibility of diploma)

-In-country study at Peking and Fudan universities: immersion integral and encouraged; e.g. partnership to establish, develop, and deliver **Chinese** In-country program at two of China's most prestigious universities, managed internally (framework, coordinated, supervised (on-site visits), subject to Faculty's evaluation, intended on-going improvement and taught by on-site staff)

-Research performance: major strategic goal (active and engaged research culture);

-incentivizing and reporting schemes: clearer expectations, mentoring and peer support structures, and financial support;
-evidence: **increase external and internal funding; high quality outputs; recent round of ERA; School's contributions** to: Literary Studies (Ranked 5), Linguistics (Ranked 4), Cultural Studies (Ranked 5), Communication and Media Studies (Ranked 4

-Research Support Officer: full time; tasks: School's research profile; strategic research planning; coordination activities: assist grant application; data collection management; meet with Admin. Manager for issues and strategies; meet with Faculty's and other School's RSOs for consistency and coordination; coordination of Research Day and Postgraduate Research day

-Research Advisory Committee: senior research leaders, strategic oversight advising the HoS (research capacity/use of funds), and junior researchers; major initiative: ARW; as APD advisors to junior colleagues (unique = external advisors to academics)

-Article Reading Workshop: peer support for researchers; draft read by one or two discussants with expertise; workshop to give feedback convened by senior committee member & organized around FoR codes; all **encouraged** to participate but now within the framework of Annual Academic Performance and Development (**compulsory!**)

-Postgraduate research support scheme: established by Univ.; **funding based on HDR enrolments;** competitive eligibility acc. to Univ. guidelines; \$ amount for conferences, travel, etc.; information session with other Schools; assessment: track record, quality of proposal and supervisor's statement

-Conference Travel support schemes: administered by Research Advisory Committee; annual budget allocation

- **Standard Conference Travel support scheme: objective: produce ERA-eligible research outputs**
- **Top Conference Travel scheme: objective: attendance top-level international conferences;**

-Research outputs: HERDC and other non-traditional research outputs: staff, research students and honorary research associates'

	<p>books, chapters, peer-reviewed articles, peer-reviewed published conference papers reported for the Higher Education Research Data Collection; in 5 yrs.: increase, with drop due to honorary associates not indicating Univ. affiliation; decline in conference papers vs. increase in books, chapters, articles due to mentoring for ERA-eligible publications</p> <p>-Research income and research applications: total income increase (huge increase); from ARC awards a huge increase (\$2m); Univ.'s initiative to diversify external funding sources: School's need to increase applications in non-Discovery Project/ARC; Research Advisory Committee: strategies</p> <p>-Relationship with External Organizations and Stakeholders: impressive record: in pedagogy?: distinguished exchange-study partners, e.g. teaching collaboration with UNSW; with Fudan and Beijing; German Academic Exchange Service (DAAD), representative in School; full-time teaching assistants, (specialists) funded yearly: DAAD, Spanish and Italian governments</p> <ul style="list-style-type: none"> • community engagement: visible and active with broad range of outside individuals and organizations (list 13 organizations) • donors (generous): list 6 organizations (to name a few) • ongoing research collaborations around the world <p>-3 challenges to be tackled (further alignment with priorities: efficiency + research):</p> <ul style="list-style-type: none"> - small size of some departments (issues of self-governance: funding) - programs to promote collaboration (European Studies and Buddhist Studies: far from realizing their full potential (not sustainable due to low enrolments) - balance teaching and research, adapt to research demands; further mentoring needed
<p>Strategic = benefiting national economy?</p>	<p>-Changes already tackled:</p> <p>-a name change: Arabic and Islamic studies >> Arabic Language and Cultures reflects "core academic strengths as well as strategic orientation" (in what way is this name name-change strategic? to attract students to study Arabic so that students can work in future trade deals with the Arabic world to benefit national economy?)</p> <p>-Buddhist studies (collaborating with Asian Studies, Chinese Studies and Indian Studies)</p> <p>-expansion of Chinese Studies</p> <p>-in-country study opports. for students of Chinese</p>
<p>Measuring efficiency: financial sustainability not accepted as only yardstick</p>	<p>- Departmental Sustainability: financial sustainability uneven; departmental deficits socialized: not undesirable situation; financial sustainability should not be the only yardstick</p>
<p>Commercial terminology: management culture</p>	<ul style="list-style-type: none"> • Financial support: part of University shared services; centralized (via Faculty with dotted line reporting to HoS; roles of Finance Manager (strategic advice to the HoS; interpret financial data; preparation of management reports (REVIEW REPORTS??) and financial projections; review of accounts (budget and expenditure expectations are met); liases with financial staff Faculty level; drafts of School budget and Finance Officer (assist Finance Manager; casual contracts, timesheets for central payment; expenditure projections; tax invoices; external research funds) • 6 month Educational designer (reporting line to the Admin. Manager) • Language Study Centre managed by the School (librarian with direct reporting line to Admin. Manager)

Strategic priorities for all disciplines, not only languages	<ul style="list-style-type: none"> • Serious challenges the School has gone through (as other parts of Univ.): <ul style="list-style-type: none"> -strong research identity -ensure strong enrolments -participate in Faculty's initiatives, take the lead, show the way
Proving commitment to Strategic Plans: cynicism, fear (expectation: not to be affected negatively)	<ul style="list-style-type: none"> • "We have been good Faculty citizens" by participating in Faculty's initiatives, taking the lead, showing the way
Justifying model: willingness to comply	-Reconfiguration: a widespread strategy not only in Australia but in "Australasia" (no evidence)
Self-praise (defensiveness/fear)	<ul style="list-style-type: none"> -<i>very close and fruitful</i> collaborations -The School regularly receives <i>very positive testimonies</i> from students -<i>very active</i> Pedagogy committee -a <i>successful</i> fund-raising campaign -is one of the School's <i>major successes</i> -offers a <i>very successful</i> unit -Internal organization has enhanced "<i>obvious</i> synergies between departments" -<i>unique</i> APD structure -<i>impressive record</i> of relationships -has <i>successfully</i> formed partnership with [...]; <i>remarkably successful</i> in-country study, long list of <i>distinguished</i> exchange-study partners (<i>best and most prestigious in the world</i>) -ongoing research collaborations <i>with institutions around the world</i> -strong progress, strong program, financially strong
Low status of languages	- carving out a place for language study in University priorities

CASE STUDY 4 (Go8) Year: 2010	
THEMES FROM REPORT	
Alignment with University Strategic Plans	<ul style="list-style-type: none"> -Context for the review (University Strategic Plans) 1. Language teaching: quality, structure, content and overall coherence; 2. Teaching methodologies; new technologies 3. Graduates: destinations and outcomes 4. Student satisfaction; demand and growth opportunities 5. Appropriateness (value) of Program 6. Performance and outcomes of language study; benchmarking 7. Cross-institutional
Praising	<ul style="list-style-type: none"> - In ES <ul style="list-style-type: none"> • very favourable findings and 13 recomms. to strengthen, ensure continued success and identify areas of potential growth • Teaching Staff: committed, responsive - Inaccurate praising: Repetition of positive findings about teaching staff in ES while reviewers' commendation highlights only section of French
("Unique") Challenges	<ul style="list-style-type: none"> • Highly variable language proficiencies • Two Schools: additional administrative issues + reduces cross-disciplinary exchange of innovations and initiatives • Add-on nature of diploma = increase of timetabling and administrative complexities • Lack of University data collection on graduate destinations and student satisfaction = limits assessment of quality against other univs.

	<ul style="list-style-type: none"> No explicit acknowledgment of importance of lang. study Not adequate emphasis on benefits of language study in current marketing literature
Realistic opportunities for growth	<ul style="list-style-type: none"> Extend range of offerings, to make available to international students unable to enrol due to gov. regulations (only in degrees, not diplomas) Offer a graduate diploma and a specialist Masters degree (Translation) to expand the range of potential consumers, additional niche markets (translation services) Improve marketing of program
Positive findings	<ul style="list-style-type: none"> Student satisfaction and performance = excellent Support for the program = strong Teaching staff = highly regarded Academic staff = highly motivated and committed and receptive to strategic opportunities to expand range of programs Languages staff's enthusiasm and dedication in French: strong collaboration with other Univs., active participation in conferences (local and intern.); ongoing involvement with external organizs. -Commendation: quality research, community engagement, innovative teaching Conclusion: language programs of a high calibre nationally: (Recent Internal Review results:) Chinese and Japanese ranked third and fourth in country; French excellent reputation: energetic involvement in research and conferences A small number do succeed: opportunity to study abroad
13 Recommends. (4 main topics)	<p>1. Facilitating increase of enrolments through</p> <p>a. Promotion (R1, R4, R6, R7 & R8) and cohesion (R1)</p> <p>b. Expansion</p> <p>-(R3): Offer a Bachelor of Languages and graduate coursework, specifically Masters in Translation (Chinese) and a Diploma in Languages</p> <p>-(R5): Strengthen and broaden the inclusion of language courses in all programs (in both schools where languages are located)</p> <p>-(R11): Engage other faculties with languages "more systematically and extensively"</p> <p>-(R12): Re-establish a cross-institutional language study agreement</p> <p>c. Flexibility (R2): Remove requirement of completing a semester before enrolling</p> <p>d. Benchmarking retention and attrition</p> <p>2. Making technology available and sharing "successful uses" (R9)</p> <p>3. Ensuring high levels of competency ("high achieving students" attain "competency levels aligned with international standards" (R10)</p> <p>4. Saving a program ("Revitalise and rebuild" a particular language area) (R13)</p>
Efficiency	<p>-ToR instruct: no recommendations requiring extra resources; indicate highest priorities in case resources become available. However, resources for some issues but not for others:</p> <p>For German, yes: why? To meet future demand for courses and for the University to capitalize on market advantage, thus increase competitive capability</p> <p>-Offerings and sustainability of offerings (Justification of 0 enrolments by international students)</p> <p>In Report of Review Panel historical overview of languages: long history, expansion (available lang. majors in other programs), Asian Study Centre + two Asian langs., Honors in each lang., Diploma of langs. not available to internat. students due to gov. regulations but very popular (enrolments),</p>

cross-institutional agreements (cancelled due to budgetary issues and unsustainable class sizes but helped open own Spanish program)
-Trends: increasing enrolments, decreasing staffing; great percentage of teaching on casual staff (implications: inconsistent delivery; student access to casual staff; threat to leave in search of permanent employment)

Promotion and Marketing of Language Study:

- Develop a language culture (inside and outside the Univ.) by
 1. Language reference group: synergies in teaching and research
 2. Improve commitment to language promotion at Univ. level
- 1 **Recommendation:** marketing language studies and development of marketing materials to attract students
- 4 **Actions:** enhance marketing of program
- Language teachers: greater role in teacher education/fostering languages in schools; lack of presence of languages in key programs
- 1 **Recommendation:** broadening inclusion of lang. courses
- 2 **Actions:** investigate programs relevant for langs. + strategies for strengthening them
- Univer.'s low profile in range of marketing activities.
- 1 **Recommendation and 1 Action:** Information Day

Language Pathways for Secondary Students:

- Negative impact: state policy, lack of identifiable career paths, University's lack of profile in this area
- Positive: French staff's involvement in professional development programs, teaching related organizations, organizing tutoring for high school students done by Univ. students
- Recommendations:** leadership role in language promotion: incentives for students and articulation of language pathways

Language Study across Faculties:

- Number of students enrolling in languages; cooperation of other faculties
- Issues: two administrations, workload, timetabling, provision of electives, inflexible structure of programs
- Recommendation:** Faculty find strategies to engage other faculties with langs.; more flexibility: study part of program in another country

Study Abroad and Exchange Programs:

- Popularity and good organization; unavailability to students with tight budgets

-**Recommendation:** investigate funding assistance; search financial assistance

Program Delivery and Information Technology Support:

- Reduced availability by Asian lang. students (Confucius I. partly funded lab renovation); find **affordable** and **appropriate** software and technology; staff training; learning enhancing technology uses from French section; **technology should not be used as a substitute for contact hours**; traditional face to face teaching style; difficulties in developing new approaches
- Recommendation:** most suitable technologies for lang. teaching; adequate and ongoing IT support; share successful uses

Language Teaching on a Cross-Institutional Basis:

- Continuing enrolments from another univ.: considerable interest in re-establishing a former cross-institutional arrangement; rewriting,

	<p>ensuring minimum class sizes, financial agreements, resource sharing</p> <p>-Recommendation: explore potential of re-establishing program: consider administrative barriers, detail financial arrangements</p>
<p>Efficiency with willingness to compliance</p>	<p>Diploma in Languages:</p> <ul style="list-style-type: none"> • Advantages: length of study shortened, support by all faculties (added value to prof. degrees) • Limitations: limited access, lack of cohesion, and sufficient administrative structure. Need of program coordinator for cohesion: course advice, promotion and marketing, oversight of enrolments, retention, tracking graduate employment outcomes and satisfaction, etc.; assumed result: cross lang. engagement, increased dialogue <p>-Recommendations: strengthen cohesion and promotion of program; enhance promotion by removing requirement of study lang. after one semester of study</p> <p>Bachelor of Languages:</p> <ul style="list-style-type: none"> • Advantages: potential instrument for broadening existing degrees due to limitations of Diploma (internat. students cannot enrol); significant interest in enrolling in electives (marketing potential for intern. Students); broader and richer experience; enhance research possibilities; broadening language study beyond language instruction; higher marketing pull than existing through Bachelor of Arts • Postgraduate degrees: Graduate Diploma in langs. and Masters in Translation to attract teachers and former graduates and international students <p>-Recommendation: explore potential offering of degree and postgraduate programs</p> <p>-Action: study feasibility</p>
<p>Justifying ‘excessive’ reviewers’ recommendation (fear of retaliation?/ servitude?)</p>	<p>-Separation of languages in 2 Schools: reviewers (unified School beneficial, but beyond its remit)</p> <p>-Recommending commitment to acknowledge explicitly importance of language study: reviewers ventured to highlight finding at the risk of exceeding their remit</p>
<p>Challenges</p>	<p>-In workload model</p> <ul style="list-style-type: none"> • Workload model: lang. staff many advantages in being in their particular School; workload model issues: <ul style="list-style-type: none"> -School 1: priority to research and engagement and not teaching with implications: conflict with increasing enrolments; research and engagement = vital for ongoing development of lang. educ. -School 2: extra contact hours not recognized; enormous value in study abroad (consolidation of exchange programs) <p>-Structural restrictions</p> <ul style="list-style-type: none"> • Importance of teaching langs. in cultural context: obstructed by new undergraduate curriculum structure (consequence: students complete only language component) <p>- Retention at advanced levels</p> <ul style="list-style-type: none"> • high enrolments in beginners (demand for more and bigger classes) diminished at later levels with “serious impact” <p>-Insufficient strategic commitment to language study</p> <ul style="list-style-type: none"> • Little evidence of a strategic commitment to langs. vs. ample opportunities to study langs.; dangerous and anomalous lack of explicit acknowledgment and omission of import. of langs. from University Strategic Plan

	<ul style="list-style-type: none"> • Cut off lang. programs <p>- external contextual hindrance + internal structural restrictions to gain proficiency</p> <ul style="list-style-type: none"> • Students' limited opportunities to gain proficiency: nothing to do with quality of teaching but with <ol style="list-style-type: none"> 1. Australia's mono-linguistic culture; schools' spasmodic teaching of languages; shortcomings of Australia's education system 2. teachers' challenge to deal with complex mix of abilities; increasing class sizes; staff reductions; adequacy and access of teaching spaces; structural changes to courses and programs; timetabling problems, all of which result in reduction of contact hours + poor retention beyond Level 1 <p>- Staffing of German threatens future viability</p>
<p>Fear of recommending against Strategic Plans</p>	<p>-Conclusion: despite challenges: the Univ. provides broad range of lang. courses + two commendations:</p> <ol style="list-style-type: none"> 1. suite, high quality of, contribution to the lang. culture of Univ. of language studies 2. staff from French: outstanding commitment to high quality research, innovative teaching, high level of community engagement
<p>Commercial/bureaucratic /management solutions/practice (FOR DOC ANALYSIS)</p>	<p>-Appendix C: Documentation provided to the review panel: Strategic / commercial in orientation = 7 categories:</p> <ol style="list-style-type: none"> 1. Introduction: interview schedule, ToR, RP's biographies 2. Strategic Information: University Top Level Organizational Structure; University's Strategic 4 yr. plan; Faculty's two yr. Business Plan 3. Public Information: public documents: brochures, printed website material 4. Program Rules: academic 5. Comparisons: summary of benchmarking with 7 Australian universities 6. Faculty Report: School submission review 7. Statistics: 4 sets of statistical data <p>-Overview of Annual University Program Performance Reports + University's submissions to the DEEWR student data collection</p> <p>-Student's course experience questionnaires</p> <p>-graduate destinations surveys</p> <p>-variety from one campus: enrolments, load, gender, etc.</p> <p>-Implementation Plan</p> <p>-Responsibility for appropriate progress of implementation plan and ensure that it is done on behalf of university (business entity, not specific people = impersonal): Quality and Reviews' office</p> <p>-Actions: develop benchmarking protocols; study feasibility of adopting international standards</p> <p>-more vigorous promotion</p> <p>-Actions: develop promotion strategies</p> <p>-Actions: appoint lang. representatives in Faculty's IT committee; annual forum for sharing practices</p> <p>-Action: needs analysis; cost-benefit analysis, analysis of options for flexible delivery</p> <p>-Actions: consider current incentive and find other ones; underscore information of interest for high school students on website</p> <p>-Actions: strategies for engagement with other faculties; study needs and benefits</p> <p>-Action: Appoint program coordinator; study requirement and implications of removal</p> <p>-Action: develop staffing renewal plan for this department</p>

	- Recommendation: regular benchmarking including retention and attrition ; support high achieving students by aligning with international standards
THEMES FROM INTERVIEW DATA	
ToR very clear but constraining	- “they did not want us er to er say certain things but we said them anyway”
Excised recommendations: Powerful authoritarian upper management	-“because they were considered not to be within the terms of reference of the” -“We said that it would be very helpful to the health of languages at the (U3) if, if there were to be in the university’s statements about its purposes some mention of languages as being useful, significant, important [at the university.” -“we were not there to give the vice chancellor advice on what his statements and policies should be” -The removal of these recommendation may not be the only piece of information that is excised from this report. Interviewee (C) comments about “Er we, we made some remarks about, about leadership and the need for leadership because er there were a number of programs that had no real, what we could call, senior leadership, no associate professor or—”
Improvement	- “now since then um there’s, there is a level E. ((laugh))” which is qualified as “Actually, that’s an improvement what was the case before hand.”
“Good process”	- Opportunity to meet: C63. [...] Now, I think that this was er a good process er the er— we had a chance to meet with our (U3) colleague before er the things started, the interview started. However, why should this action be considered positively? Is it important? Should this happen at all? (French section the most praised in report) - Like-minded reviewing group: “we found the process, all of us I think, quite congenial” E.g. internal member of the university in the panel “who had no connection with languages at all, [...] “in the end she was as vigorous in her support of the recommendations that we made er as anybody” - Positive unintended consequences: learning internal member of panel “who had no connection with languages at all, she learnt a huge amount from er being there, from participating in the process” - Opportunity to contextualize findings C75. [...] So um now, in the, in the (U3) thing, the er figures we were presented with, we were able to test against the er people who we interviewed um and they had— - Diversity of informants: “we got to talk to a lot of er a lot of different people including er some people from er neighbouring universities with whom (U3), (U3) had some agreements in relation to sharing er language teaching”. - External language representation: “you had representation, um external representation of experienced** — external representation of the languages areas. I think that’s, that’s very important” E.g. interviewee (C) contrasts this review process with one which lacked language representation which incidentally is CS6: C75. [...] For instance, at the (U4) where, I mean, in this review that they did of the (L6) program, there was no representation of language at all, and I don’t think anyone external or they might’ve been someone external but it was not somebody to do with language [...]

	<p>-Review described superficially but positive aspect of review based on “strengthening” purpose of review associated with expansion of program (perceived as exceptional): K72. [...] my recollection— when I had to go through the documents was that overall it was a very positive review, with a number of recommendations to try and strengthen languages, including new sort of degree structures and majors, I think. Er—</p> <p>-Good level of program staff ‘s engagement with review (perceived as exceptional): when interviewee (C) is asked about the level of engagement of program staff in the review (School Submission Review?): C75. [...] it turned out that they were the people who’d put the figures together</p> <p>C97. Chances are, I think it depends hugely on the, on the thing um I mean, in (U3), the case that I gave you it was very clear that the members of staff er were aware of the, of the terms of reference of the review, and may even have participated in the formulating. I think in a lot of cases that doesn’t happen.</p>
<p>Reviewers</p>	<p>-Background in languages more valued than expertise in evaluation ** The term “experienced” must be read with care. While the meaning is not fully expressed as the transcription symbol indicates (unfinished sentence), in this context it may mean ‘experience in the languages profession’ since the interviewee explains at another point in the interview that “I don’t consider myself an expert evaluator”</p> <p>-Expert evaluator perceived as unrealistic and as a problem (C) expands on notion of expert evaluator by explaining personal view revealing that there is not such a thing as an expert evaluator and furthermore highlights that “expert evaluators” may be one of the problems with reviews.</p> <p>-Due perhaps to the fact that reviewing in an unpaid job? Three different interviewees show how reviewers are not remunerated for ‘serving” on Review Panels and in (T)’s case, the fact that it is “unpaid” is contrasted with the “enormous amount of work” thus signalling a perception unfairness.</p> <p>O26. [...] and, you know, these are senior people often, you know, they don’t normally charge or get paid honoraria for doing it. You just do it as part of your work um and, in no case— [...].</p> <p>R34. Well, that’s where,] yeah, no, it’s not remunerated.</p> <p>T56. Er also, we need to remember that the job of sitting in one of these panels is very much a freelance kind of job. You don’t get a remuneration for it, and I can tell you that the number of hours that are going to sit in one of these panels, and the level of, well, personal involvement that goes with it, sometimes er intense interpersonal relations, is very high, I mean, is an enormous amount of work, which is unpaid. [...]</p> <p>-Participation in review panels = opportunity for commissioners to attract talents to achieve priority goals + for academics to make progress in career -“the other er reviewer has since become a high level er person at the (U3). He was, he was poached ((laugh)) from a, from another Group of 8 university”</p>
<p>University priorities associated with building a</p>	<p>C65. Yes, they wanted to know whether the— what their processes were like, what their, what their, what their programs were like er in comparison to others around the nation, I mean, this is a part of the university’s hefts to position itself (xx) as a, as a research—</p>

reputation and staying competitive	
Student outcomes and program outcomes mentioned not to do with proficiency outcomes	-lack of proficiency outcomes “which I have never seen any department do”
Double report (Reviewers/Internal)	-“was excised from the final report”

CASE STUDY 5 (Go8)	
Year: 2010	
THEMES FROM REPORT	
In this report, requested areas to be studied associated with alignment of evaluand with Strategic Plans (Review questions) and improvements are generally associated with aligning with broad priorities with emphasis on enrolments and research; delivery = range of courses, broad questions e.g. contribution to undergraduate education, to internationalization	
Priorities	“education, research and outreach”
Efficiency	<p>Enrolments</p> <ul style="list-style-type: none"> -Apply more widely lectures taught in English model and tap markets for students in wider university -Reconfigure BA: 4 yrs., with language study and Linguistics, with mandatory study abroad -Implement pedagogical strategy: modularise language courses into four skills and admit students at differing levels -review of Honours Programs to offer: option of writing Hon Thesis in English -Urgent strategies to increase enrolments in Masters of Applied Linguistics and Master of Translation -Increase HDR numbers “as a matter of priority” and identify specific target markets -collaborations with government to improve Language teaching skills in secondary education -develop Hons. Program with other Go8 Univs. -Ways to maximize student numbers in taught Masters programs -Ways to increase student numbers in higher degree by research <p>Research</p> <ul style="list-style-type: none"> -Appointments for phonetics; Indigenous Languages; Language, Literature and culture (w/ strong language-centred research profile) - Indigenous langs. collaboration with a specific Univ. -“as a matter of priority” improve research performance: increase research active staff & success in external grant applications -As part of curriculum review, embed research in undergraduate curriculum -School’s research performance and ways to enhance it <p>Outreach</p> <ul style="list-style-type: none"> -maintain relationships with embassies (in relation to the year abroad program) <p>-Organization and structure:</p> <ul style="list-style-type: none"> -lack of senior academic positions, partic. no professional appointments in lang. in 10 yrs. • Administrative support: -Staff thinly stretched; no full time administrative staff; insufficient support for School size and outreach actv. • Facilities and staffing:

	<p>-retirements not filled in to teach and research core fields of expertise: immediate staffing appointments for program coherence and ensure reputation national and international in linguistics particularly</p> <p>-lack of well-equipped spaces for teaching langs.; current facilities inferior to other Go8 univ.'s and inapprop. for 21 cent. Lang. teaching & for a Univ. aspiring to be "the leading language teaching provider"; refurbishing and redesigning plans or find alternative teaching spaces</p>
<p>Alignment with Strategic Plans and priorities associated with efficiency</p>	<p>-In review questions:</p> <ul style="list-style-type: none"> -Articulation of School's courses and majors with structure of Univ.'s curriculum -Articulation of School's Masters programs with others in the College -Articulation of educational and research programs with College's strategic plan <p>-In Recommendation (range and delivery of courses): build in flexibility in line with University's educational planning</p> <p>-In findings: Outreach "entirely consistent with stated core business of the University" (e.g. Linguistics Olympiad: very successful in attracting students; should continue)</p> <p><u>-Organizational and Governance Structures:</u></p> <p>-Questions: organizational and governance structures; appropriateness of administrative support; appropriateness of current facilities and staffing to fulfil education and research mission</p> <p>-Needs improvement: Appointments:</p> <ul style="list-style-type: none"> - fulfil School's strategic mission: 2 level E appointments -Managerial: School Manager: supervisory responsibility, budget and strategic planning skills, build relationships with external partners - support HoS: reconfigure 1 administrative role -Academic: Level B/C junior appointment <p><u>-Disciplinary Balance</u> (= linkages = collaboration)</p> <p>-Questions:</p> <ul style="list-style-type: none"> -enhancement of linkages between School and sub-disciplines -development of European languages and cultures, linguistics and applied linguistics to enhance undergraduate and graduate experience -linkages that contribute to this enhancement <p>-Findings:</p> <p>-lack of collaboration/integration in curriculum planning + research between Langs., Lingus. and Applied Lings.</p> <p>-Needs improvement:</p> <ul style="list-style-type: none"> -greater clarity of: learning outcomes, level of expected students' lang. acquisition, nexus between lang. and culture; nexus between teaching and research -staff development program: contemporary international debates about language/culture nexus -reposition of language disciplines as vocational specialists investigating the "language-making capacities of culture" and vice-versa; result: greater collaboration -greater opportunities for staff professional development (e.g. apply for grants to visit other instituts. to learn ab. Curriculum innovation and best practice) <p>• Linkages within the College</p> <p>-Positive findings:</p> <ul style="list-style-type: none"> -Good linkages; encouraged to develop links further; important contribution to the BA (Intern. Relations); range of very successful thematic courses (Italian for opera singers and opera lovers), appealing to students with or without fluency; structure them for widening access (College and University)

-Needs improvement:
 -Balance between lang. **specialist courses and others for wider audiences**
 -greater engagement of staff from other areas in **thematic courses** (e.g. film, history); **Latin American Studies and Asia Pacific should influence foci of lang. programs**; work with School of Culture Inquiry to **develop initiatives** in comparative literature

• **Linkages with College of Asia and the Pacific**

-Needs improvement:
 -**greater collaboration at postgraduate level (A. Linguistics and Translation)**; collaboration with Asia and the Pacific to **develop study overseas program**

-Range and Delivery of undergraduate studies

-**Questions: contribution to undergraduate education** in Australia and **internationalization** of education; appropriateness of **range of courses** available; Degree to which courses and majors prepare students for **Postgraduate coursework** and Higher degree by research

-**Positive findings:** Good range of programs and courses; the University's distinctiveness in **expertise in Indigenous Langs.** (staffing appointment highest priority + great **alliance potential** with a specific University)

-Needs improvement:

-**BA of Languages:** developed and re-profiled modelling that of the College of Asia and the Pacific **to ensure “administrative efficiencies” and “congruence of practice”** (congruence with what?):
 -stronger focus on Lang. and Linguistics;
 -build in flexibility in line with University's educational planning;
 -include a major in a lang. and a second major in Ling. or a second lang.
 -Length: 4 yrs.
 -**Include overseas study** (a year: “badged” A year in Europe or in Latin America) with option of shorter stay provided

- For “the period abroad”: **Look for funding** and consider issues such as access, equity, cost of living of Europe in comparison to Asia
 --Cross-college coordination (studying two langs.)

-Various levels of proficiency among Heritage speakers: **modularization of lang.** (search: modularization associated with a flexible language curriculum) four skills; entry assessment for correct placement

-**Desirable:** offer **Portuguese and Dutch** (subject to resources); **market unlikely large and therefore flexible modes of delivery** (e.g. online; Dutch one yr. program and Portuguese “in alliance with” Spanish)

• **Honours:**

-**Needs improvement: low enrolments** in European langs. (general **national trend**): very difficult to recruit students from within the School; concern about **thesis written in target lang.** rather than English; School's belief: elite status vs. panel: disciplinary mastery separated from language competency; panel recommends **option to write in English** and given national trend panel suggests the university take leadership role in **developing national collaborative Honors Program** limited to coursework component, not thesis component.
 -some theses completed in subjects which require interdisciplinary expertise (e.g. Spanish + political science): supervision by members of various disciplines possible **without expecting students' expert knowledge** in those areas (e.g. political science methodologies)

	<p>-Students: no formal coursework in some programs: common assessment framework across languages developed</p> <p>-Panel's suggestion: substitute Honours for a "Master of Research" as C. of Asia and the Pacific model (= more income) with year abroad accredited to enter "M. by Research" (= enrolments)</p> <p>-Capacity for further development of taught Masters programs (= post-graduate offerings e.g. Masters)</p> <p>-Questions: Opportunities for development (= a priority: more income)</p> <p>-Positive findings: development and anticipated success (= income) of Master of A.L. and Master of Translation jointly with CAP (C. of Asia and the Pacific); consider expanding translation to Spanish and Russian but first explore market demand</p> <p>-Needs improvement:</p> <p>-little progress in M. of Linguistics and M. of Languages and Culture (= interest in higher degrees: income): urgency to tackle issues: nexus between languages and culture and leadership</p> <p>-explore collaboration at graduate level with School of Music and School of Cultural Inquiry (e.g. comparative literature) (= increasing enrolments in higher degrees efficiently)</p> <p>-Higher degree by research (= students produce research that count as publication outputs)</p> <p>-Questions: control and alignment: Level of success of School's research student; Distribution of student numbers across disciplines; training programs for doctoral students; introduction of a coursework component to doctoral programs</p> <p>-Needs improvement: underperformance of HDR student enrolments "due to a mix of factors":</p> <p>-lack of senior academic leadership</p> <p>-research dimension not incorporated in undergraduate languages</p> <p>-disconnect between Honors and doctoral programs</p> <p>-limited scholarships, partic. for internat. students</p> <p>-Priority: attract students into HDR (e.g. identify markets in Asia and Latin American's students with access to local funding for studying abroad)</p> <p>-Outreach to the community</p> <p>Questions: Outreach activities, level of success (= attracting income), measurement of success (only question in whole study requiring how success is measured, knowing that success = income)</p> <p>-Positive findings:</p> <p>-in School's submission = "considerable extent of outreach activities" "entirely consistent with stated core business of the University" (e.g. Linguistics Olympiad: very successful in attracting students; should continue)</p> <p>-obvious capacity and opportunity to collaborate with government in relation to language planning and policy: collaborate to upgrade language teaching skills in secondary schools</p> <p>-embassies show great interest in languages: maintain and develop further connections with embassies for funding of "a Year in [...]" program</p> <p>-National and International research performance:</p>
--	---

	<p>-Questions: Comparisons across disciplines and languages; Level of integration of educational and research programs; Mentoring for academic staff</p> <p>-Positive findings: long international reputation of research excellence in Linguistics; for stellar position not to be lost: staff renewal and renewed leadership</p> <p>-Needs improvement: -languages: less strong, substantial improvement, lack of dynamic research culture, lack of clarity about areas of research excellence possible due to lack of senior appointments therefore mentoring -undergraduate education needs to be delivered by leaders in research: not based on researchers' areas but on interesting questions in the field; clarification of relationship between research expertise areas and the field can be achieved through review of curriculum</p>
Improvements	<p>-Improvements: -redesign and refurbishment of interactive teaching spaces and lang. classes timetabled in technologically equipped rooms</p> <p>-interdisciplinary thesis should have combined Honours program; joint supervision for combined Honours; develop formal coursework for Hons.</p> <p>-stronger leadership/internal School structure for greater integration of constituent parts; effective School Executive; succession planning (more opportunities for junior staff to acquire leadership skills)</p> <p>- Improvement associated with strategic priorities (beneficial for program and in alignment with priorities) -Undertake curriculum review to: identify learning outcomes of programs; define students' level of acquisition and benchmark w/ international standards; develop more sophisticated theoretical nexus between language and culture</p>
Focus of school's description on structural changes	<ul style="list-style-type: none"> • Dates of additions, creations, establishments, disestablishments, restructuring, appointments, abolitions, no replacements, reductions <p>-Willingness to compliance?</p> <ul style="list-style-type: none"> • "less material impact" in comparison to other Schools by most recent change (from Faculty of Arts to College of Arts and Social Sciences) • only change of note: incorporation of a Dictionary Centre
Overpraising (= defensive)	<ul style="list-style-type: none"> • Purpose of review: "to raise the education and research profile of the school with a view <i>to taking its rightful place</i>" • The Bachelor of Languages is <i>logically</i> the flagship program of the School.
THEMES FROM INTERVIEW DATA	
Triggers	<ul style="list-style-type: none"> • A concern: (M): -"the School and the Faculty had a reason to be worried about the languages" - losing Honours students • Periodic reviews: - (I): "embedded in the rules" • Subversion: the internal cycle of reviews is used as a pretext to make changes such as replacing a particular individual resisting changes and replacing the position with someone else, who in fact was a member of the Review Panel: <p>I13. [...] And not only that, they also brought someone from outside as a professor to put it as the new head of school in the next year but the review was conducted before this person took the place but she was part of the committee, she was part of the review.</p> <p>Q14. Of the panel. I15. Of the panel [...]</p>

	<ul style="list-style-type: none"> •(I): Targeted review: the commissioning of the review was taken advantage of in order to make a personnel change, due to a “battle”, a “clash” between a dean and the HoS, to in turn push changes further
Program context	<ul style="list-style-type: none"> •(M): Disadvantaged: <ol style="list-style-type: none"> 1. push to make it more cost effective 2. lack of understanding by commissioners about language teaching and learning, 3. program’s unfavourable position in the university 4. pressure and lack of institutional support
Review purpose	<ul style="list-style-type: none"> • Set priorities to align evaluand with (I): ERA (Excellence in Research for Australia) to push focus on research and secure ranking positions: <p>I60. [...] from having someone like (P23), who was a very experienced player in the field, to have someone like (P24), who is someone who was able to bring er one university, that was (U17), from nothing into the— all the research charts that are used [and so basically—</p> <p>I60. [...] the review also was gear very much to align our school with the needs to be very successful in the incoming ERA, the Excellence for Research in Australia.</p> <p>which required to</p> <p>I60. [...] produce research as a priority and not to put so much time into teaching.</p> •(M): To increase honours enrolments: <p>M12. [...] to increase um their high degrees and their honour students, and make more sense of what they were trying to do.</p> • Compliance with policy to monitor how areas of evaluand under review are managed • to align School with broader university goals with a focus on funding. (I): <ul style="list-style-type: none"> - “an imperialistic move to change the er practices that were maintaining the quality of language teaching against practices that were more— perceived more as more commercial or more er towards er sort of aligning our university with practices at other universities, including Group of 8 universities, but which people perceive them as not, not good. - to impose change that was not welcome. - it was essentially to impose a lot of cuts and changes • to push programs to improve but in conditions that are not conducive to good practice and which may also lead to subversive practices: <p>I66. Learn to live with and, and it’s a very difficult thing because, I mean, at a lot of levels er what we have experience is, of course, I mean, ‘reduce number of tutors er and, and er you have to cut corners here and there, what is the least worse outcome in a bad outcome’ and all this kind of things, but it’s, I mean, they are trying to fulfill the—</p> <p>Q67. They have an agenda, is that right?</p> <p>I68. Well, of course, they have their agenda, they have their agenda of, of having this, and this, and this, and this parameter to be improved and it doesn’t matter how it improves, it has to be improved. And so er at the level of the language teaching, we have been able to maintain a lot of the quality but using a lot of</p>

	<p>ingenuity and a lot of imagination, simply because, I mean, the restrictions are big.</p> <p>I132. Oh, that was,] that was very difficult, yeah. That were the things that were difficult for them to understand, and so we got— I mean, and, and, and there was, of course, also the issue of the overarching issues. I mean, the, the, the overarching sort of pushes: ‘You had to comply with the ERA, you have to be prepared for this, you have to do that and that and that’. And so, at some level, the problem is that, I mean, you are forced to work er under constraints that are completely unsuitable for the language teaching environment.</p> <ul style="list-style-type: none"> •Additional priorities: Curriculum change: <p>I31. [...] together with the review, there was an overarching change of the curriculum in all the university but specially in the faculty of arts that was the biggest faculty in the, in the university [...]</p> <ul style="list-style-type: none"> •Institutional priorities tied to national economy priorities with a particular ideological slant: <p>I190. [...] If we had a system— if we were in a country that really see education as a value and not as a tool to economic advancement, because this is how it’s perceived, and how it’s argued, this is what they are arguing now: The private benefit is more important than the public benefit, I mean, we are coming, I mean, I’m coming [to the country—</p> <p>Q191. So is that a process] of priva[tization of the university? I192. Yes, yes, yes.]</p> <p>-which is turning universities into “for-profit” organizations concentrating efforts on financial issues and on control:</p> <p>I192. [...] I mean, when I have to explain to my— I mean, in (COU1), I was, I mean, in a very privilege university. This was a very, very privileged place, I mean, I was a mathematician, I was in a very privileged place that had funding, external funding. So we had a library that when I went abroad, I realized it was extremely good. It was funded, I don’t know now, but, I mean, at the time, it was funded by grants, external grants and we had, I mean, a flow of visitants, most of the students went to study abroad, and this kind of thing, so it was not the normal university but when I had to explain to them our practices and the kind— they say ‘this is even worse than the private, only for profit universities’ that they see in the same city.</p> <p>Q193. When you explain to them in (COU1) about the Australian model.</p> <p>I194. Yeah, about the Australian system, about how— what things are happening here, they say ‘well, this is even worse than (Uvii) that it’s basically a (xxx) to get students. That’s, that’s the contrast, I mean, but it’s because, I mean, at least where I was there was a perception that academic excellence was the goal, and the funding was geared towards that. You didn’t have to tick all the boxes at the performance reviews in which it shows that you have done enough outreach, enough research and so, first, enough research, then, your teaching is not defective according to how students evaluate you. It doesn’t matter—</p>
Review focus	<ul style="list-style-type: none"> •(M) Rotating or not and whether specifically or not to do with the languages, reviews tend to focus on administrative issues (student numbers, retention rates, etc.) and not on the <u>work</u> of a language program
Reviewers	<ul style="list-style-type: none"> •Range and number of reviewers: -(I): expected (“of course”) -(M): high number due to size of evaluand

	<ul style="list-style-type: none"> ● Insufficient background in languages (M): -“not a concentration of language specialists”; “So there might have been one, one linguist and one er applied linguist but he was also a dean in another university” (managerial role neutralizes/invalidates academic expertise); “I was the person who knew most [...]”; “I was chosen for [holding a prestigious honorific title] ● panel worked “tremendously well”, very cooperatively: panel members did not understand but “they came along with me”, “particularly” “the dean” and “a higher dean like a PVC”, “who had chosen me by all accounts” ● inadequate selection: only language teacher in the panel. (I): -“it was not adequate”: non vocational, research oriented, limited practice in application which resulted in “a lot of misunderstandings”, “a very dramatic event” ● Panel’s lack of knowledge about language teaching contributed to ignoring evidence: (I): “they would not listen” ● consultation process for selection of reviewers and initial inadequate selection (I): -“a very confusing affair”: “not wide enough”, “some of them [...] appointed without consultation”; consultation of a “language teacher” done in a rush by the head of School when staff complained about absence of such background, ● Role of reviewers. (M): -Expert assessor (making positive suggestions), advocator of languages (get committee on side with the languages) and educator to other members of RP (clarify doubts to panel members about program) -due to unfavourable program’s background = RP, a threat -(I): “To comply with terms of reference “ (M) (External reviewer with background in languages) with preconception about the program: (M): “we all knew there was a problem and we all knew it in different ways” M8. [...] It has to be said, this is my private view, that (U14) languages has always been very weak, [...] I19. [...] that we were weak disciplinary, disciplinary weak, which was the preconception of this woman anyway. Role: -RP’s role: To use arguments based on biased/prejudiced perspectives about the program to support and justify the “pushing” of changes -RP = Informants for anonymous writer of report. Anonymous writer of report assumes role of reviewer (meta-reviewer?) who re-reviews stakeholders groups’ perspectives in which the Review Panels’ are treated as one more set of perspectives -(I): Role was not but it should be: to make people understand the negative effect of impositions, partic. cost-saving ones, on language education
<p>Methodologies</p>	<ul style="list-style-type: none"> ● “feedback from staff”; “anonymous” collected by (I) to present to the review; enrolment statistics, exchange, gave it to the HoS “who gave it to them” ● (I): Irregularities.

	<p>- staff surprised to learn about only some submissions were made public; people from other areas were invited to make submissions (= some decisions about procedures are perceived as ad hoc therefore trigger suspicion which in turn reveals a lack of knowledge by staff about the organization of the review process)</p>
Commissioners	<ul style="list-style-type: none"> • (I): prejudiced: I116. [...] when it happens that you have someone like we have, that have a lot of prejudice, didn't— was not permeable to socialization
Consequences	<ul style="list-style-type: none"> • Negative changes. (I): -as a consequence there were a lot of different practices that had to be changed and we had to maintain things with different, with different strategies - we were restricted and casual into doing lectures and doing a lot things that were not conducive to language teaching. - as a consequence of the review, [...] we lost the position of having in (D1) a compulsory major. And so this— I mean, today the only program that has a compulsory major in the language at (U14) is the bachelor of languages and I'm the (PC1) of it and I usually joke that er I still have to fight with administrators that will, some day, come to me and say that 'I don't have to have a compulsory language major in the bachelor of language'. ((laugh)) • Negative consequences: • Pain, stress, conflict: I116. [...] a push to impose [...] that result on a lot of pain because essentially language teachers will not relinquish the issue of teaching the language well, but they will have a lot of extra impositions that they will have to comply with. I52. [...] So it was a repeat of that, here, but with people that were— had a long sort of way of doing things, that would not sort of be flexible enough to adapt. So this has created [a lot of stress]. I90. [...] So this was— there were a lot of, of, of levels of conflict because there was the level of conflict by just having a review that was perceived as an imposition, plus the level within the school on deciding strategies to face this, and then the key • (M): recommendations "taken very badly"; "total paranoia"; negative emotional reaction (defensiveness and hurt) • growing animosity between managers (commissioners /organizers) and staff: I126. [...] so this was a positive of the review. But it required a lot of militancy in the sense that, I mean, you have to show them how things were bad, and go against the dean and the, and the education dean, who were feeling ashamed that you were telling people of the outside members of the panel how bad things were because of their diligence. I128. So you didn't— I mean, we won that. The recommendation was to upgrade and to do— but, of course, this was an immense cause internally. • Post-review workshops given by external reviewer with language background (never carried out) -Program's interpretation or double report-writing? M16. [...] So, he told me, he was here for conference and he said 'let's have coffee, I think I should explain something to you'. So we had coffee and he said to me they think this, that and the other, meaning that he thinks that of course, and I said 'look, (P25), I'm not coming to do workshops under those circumstances', I said 'I didn't say that. That's not what the report said, this is not'

	<ul style="list-style-type: none"> • Increasing mistrust in colleagues from within the discipline: (the most awkward one, the most difficult one [review experience]”; the head of the department “he was the most defensive of all” • Disheartened: “I don’t care”; “they can think what they want but it’s a real shame” • (M): Ego, prejudice: emotional passions vs rational judgment assessment about review may ensue from partial and personal perspective influenced by a biased pre-review perception about the program and by (M)’s expertise questioned • (M)’s uncontrolled passion of the ego leading to denial and contradictions: <ul style="list-style-type: none"> -denial: “I didn’t say that” (against the bonding of language and culture) -contradiction: recognition of the cause of the concern: that recommending to raise student numbers by suppressing the requirement to write in target lang. could mean increasing chances of teaching culture without the L2, “free standing”, but without recognizing this implication due to passion of the ego: “total paranoia”; “they wanted for us to tell that they were very good at their jobs and they needed more money to do it” -denial of political use of review (to push for changes to ensure priorities are tackled regardless of how it affects the standing of the discipline) and blame negative response to recommendations to staff’s misinterpretation of report, program’s staff “paranoia” and self-interests involving a refusal to changes
--	---

CASE STUDY 6 (Non-Go8)

Year: 2012

THEMES FROM REPORT

Biased review ToR and process	<p>-commissioners’ interest to safeguard educational continuation of students = a source of bias indicating commissioner’s inclination to use review to implement pre-review decision = close down program</p> <p>“[...] including alternative options for students wishing to study Japanese language in the ACT”</p>
Proving false triggers and assumptions	<ul style="list-style-type: none"> • Program description: <ul style="list-style-type: none"> a) “attracting” students both local and international b) alignment with both University and International policies c) responsiveness and adaptability to changes “over a short time” d) active outreach engagement both nationally and internationally e) proactive engagement with marketing f) “main driver” for exchange • Summary of findings • Key facts <ul style="list-style-type: none"> (a) responded to the 16 recommendations from past reviews (b) implemented “numerous” actions to increase enrolments (c) reduced weekly classes (d) increased enrolments “threefold over eight years” (e) financially viable since not only have teaching costs nor surpassed student fees but the program has contributed a profit • Concluding Statements <ul style="list-style-type: none"> 1. Beyond finances <ul style="list-style-type: none"> a. Program’s contributions -Program’s success in aligning program with both the Univ.’s as well as the Federal government’s strategic goals

	<p>- Financial criteria should never be the only type of criteria to make a judgment about a program but also “results in achieving strategic goals, alignment with key policies, contribution to campus culture, and future developments”</p> <p>2. False review triggers and assumptions</p> <p>-Conclusions:</p> <p>-Concern: misinformation in the triggers of the Review Notice</p> <p>-Compliance with Review Notice</p> <ul style="list-style-type: none"> •-List of RN’s triggers commented` : denying assumptions as demonstrated in Results Ch., concluding: “We have proven that the major triggers for the review were based on misinformation, and the premises on which the review questions are based are not accurate.”
<p>Relativisation to change focus on to evidence presented in report</p>	<ul style="list-style-type: none"> •softening of fourth trigger in the self-review report: “Community concern about the proposed closure has focused on issues of educational and academic integrity, and procedural fairness; all of which are warranted” used to focus commissioner’s attention on other three triggers for which ample actual data presented in Results chapter
<p>Vulnerability of program/High stakes of review</p>	<ul style="list-style-type: none"> • Evaluand not a language degree or diploma but major in Japanese since “As of 2012, there is no Japanese or Modern Languages degree or diploma offered” at this university. • call to Review Panel revealing high stakes decision (and lack of confidence in neutrality of RP?) “We trust that the Review Panel will [...] and “urge the panel [...]” to take into consideration the analysis, results and recommendations “in this paper” and their hope that their decisions “will be fair ones based on data, sound analysis, and coherent policy” • Closing down Jap. program = losing students • Recommendation: retention of program 1. Maintain and expand program: justified with description (from ES) highlighting program strengths • Appendices: A2. Commendation from the most recent review (2010) A9. High School supporting letter • 12. Supporting letters from alumni and current students of the Jap. language area (48 pages)
<p>Defensiveness against measuring program against financial viability only</p>	<ul style="list-style-type: none"> • Program description -positive impact “beyond student numbers and status as a service area”: a) “attracting” students both local and international b) alignment with both University and International policies c) responsiveness and adaptability to changes “over a short time” d) active outreach engagement both nationally and internationally e) proactive engagement with marketing f) “main driver” for exchange
<p>low status, pressures, confidence in program’s fine condition and compliance with U.’s guidelines</p>	<ul style="list-style-type: none"> • Comments or Actions • Proposed Initiatives: to further enhance financial viability: a) Reintroduction of advanced units (4 actions) b) new course offerings (4 actions) c) developing a liason with target language country government program (2 actions) d) review financial model for assessing internal charges (4 actions)

<p>Disproval of assumptions in Review Notice Misinformed premises in Review Notice Small enrolments don't always mean deficit</p>	<p>-Recommended immediate action: a) for it is the one to most reliably bring in a "sizeable surplus" despite its small enrolments</p> <p>• Recommendations -2 recommendations: b) reintroduction of advanced units c) expansion of program through "formal" cost done "against whole-of-university benefits"</p>
	<p>RESULTS CHAPTER:., proving assumptions in Review notice wrong,</p>
<p>Demonstration of unnecessary review</p>	<p>-program under review never the focus of past reviews -program's financial viability never the focus of past reviews</p> <p>• Student numbers: actual student numbers looked at in isolation and in comparison with units of study within Faculty; results: numbers not only a concern but tripled over 8 yrs.; current enrolments not "small"</p>
<p>Proving inappropriate/unprofessional administration practice/incompetence</p>	<p>-languages involved only in 3 of the 4 reviews mentioned -staff's resistance to changes due to management of "multiple restructures and management initiated changes over several years" done "often without meaningful consultation, apparent evidence base, effective change management" -Contradictory recommendations from one review to another -Errors in reports about previous reviews -Important difference between attrition in sequences/ electives vs. in degrees -Denial of statement in Review Notice: teaching costs not exceeded income from tuition in past 3 yrs. -Difficulty in obtaining data "after repeated requests"; made available 4 days of submission deadline with data covering only four last yrs. -Inaccuracies: significant errors in calculation; corrections and misallocations of revenues and expenses; misallocated and double-counted costs; conclusion: true surplus; "program was making a specific financial amount contrary by assertions made by the dean [...] and in the notice of this review"; -No deficit clearly demonstrated; -Recommendation: 'Overhaul and radically improve" information management systems in order to access data and make "meaningful analysis and forecasting" of student movements, alumni trajectories, course structures and financial information -Initiatives: review financial model for assessing internal charges (4 actions)</p> <p>• Modelling costs's summary demonstrates student load should not be taken on face value: 1.5 times per student; lang. lab, equipment at other Univ., not this one -Lang. teaching more viable when adjusting financial model: reflecting gov. support, extra 50% funding vs base-funding cluster which covers the extra cost; internal charge should discount this extra money. -Calculation presented: much fairer, more true to Gov. intentions internal funding model; urge Fac. To change current model into this one</p> <p>• Recommendation 2 highlights consequences of unprofessional management practices and decisions: Improve information management systems: justified with negative resulting events: attempted closure of program, staff's wasted time, unnecessary review, reputation damage</p>

<p>Efficiency (budgetary discipline and generating income)</p>	<p>-“the vast majority of” changes have been implemented “within resource and structural constraints”</p> <p>-Work (from recommendations from past reviews) taken up by program prior to recommendations but only partially achieved due to Univ. degree constraints</p> <p>-7 strategies considering: no control over degree structures</p> <p>-degree reform: reduction of mandatory subjects, increasing of electives</p> <p>-Student numbers: Close acquaintance with issue: “repeatedly pointed out to us”</p> <p>5. Winter term offering of Jap. 1A: Language and Culture: Following Faculty’s request; to increase second semester enrolments</p> <p>-Diploma in Languages: discontinued but important since it allowed anyone to take lang. study (why discontinue?: contradictory or malpractice?)</p> <p>- No degree course in Jap; less presence than ANU (less competitive? contradictory, dysfunctional practice (=incompetence) or prejudice against languages? or a mix of the two?)</p> <p>-Univ. looking to increase Chinese students in future; closure of program = popular elective no longer available; less attraction for students to come to Univ.: in program not as severely affected by limited English skills as in other areas (contradiction: incompetence, prejudice or both?)</p> <p>“Classes spread over several days” emerged in previous review; recommendation implemented with no success: classroom and staffing availability</p>
<p>Focus of report on viability</p>	<p>-Attrition, Student numbers</p> <p>-classes spread over:</p> <p>-costs against student fees:</p>
<p>Demonstrating awareness of priorities and willingness to comply with priorities and beyond (= educational quality)</p>	<p>-Program description:</p> <p>a) “attracting” students both local and international</p> <p>b) alignment with both University and International policies</p> <p>c) responsiveness and adaptability to changes “over a short time”</p> <p>d) active outreach engagement both nationally and internationally</p> <p>e) proactive engagement with marketing</p> <p>f) “main driver” for exchange</p> <p>-opportunities for growth in student numbers and enhancement</p> <p>ADDRESSING PAST REVIEWS</p> <p>Attrition: Program’s interest in monitoring attrition: study of true retention rate in major and first yr. student numbers over 8 yrs.; related research consulted; 3 conclusions: retention rate does not correlate with student numbers; yr.-long student numbers improved, yet proportionally students discontinued after first semester; financial viability increases with recruitment of single elective students</p> <p>-benchmarking: better retention</p> <p>• 7 strategies:</p> <p>1. Fostering student’s study habits through learning outcomes adjustment: encourage students to develop habits; crucial in the first yr.; adjustment of learning outcomes and generic skills applied to course content and assessment (Appendix)</p> <p>2. Mentoring system and its improvement: peer mentoring for supportive advice outside of teaching staff; research literature: peer mentoring helps retention; lessons from the past: students in most</p>

	<p>need not attending to sessions, reasons, changes applied; identification of students at risk</p> <p>3. Internship students from Japan: Japanese students training to become language teachers from internship program in partner Japanese Univ. giving extra support to students and getting experience themselves; number of students received and contributed income to the Faculty</p> <p>4. Class participation and support from Japanese exchange students: exchange students: observers and practice partners for students; formalisation of program: support for 1 yr. students; identification of students at risk; study groups with exchange students; partner univ. willing to require this of their exchange program students; restructuring contents of 3 yr. major: reduced grammatical content, increased revision and communicative activities to allow students identify if behind and allow to catch up; new set of textbooks for helping identified weakness: kanji character learning</p> <p>5. Winter term offering of Jap. 1A: Language and Culture: Following Faculty's request; to increase second semester enrolments: second chance for first sem. failures + mid-yr. students; lessons learnt from low enrolments: potential causes: insufficient marketing/low demand; marketing may increase but decided not to run: may end up serving students with course as elective which, brings revenue but does not have flow-on effect; however, willing to re run it if Faculty and Univ. view it as beneficial; actions to be taken</p> <p>6. External factors: degree reform: particular degree with an elective major; potential work with them to increase enrolments</p> <p>7. External factors: Japanese lang. learners in ACT schools: 253% increase of Jap. Students in second. Educ. in previous three years; expectation for increasing numbers and better qualified students thus for higher degrees;</p> <ul style="list-style-type: none"> • Student numbers: ongoing efforts; different strategies; willingness to maintaining financial viability "we are passionate about attracting students" <ul style="list-style-type: none"> - 6 Internal recruitment strategies: <ol style="list-style-type: none"> 1. Offering attractive unit content: constantly revising course content (students' needs/ever-changing educational environment; Appendix); seven changes of the "numerous" ones presented; high praise in past review: innovative practices (i-chat in collaboration with Jap. Univ.; real life materials through internet; Moodle for aural skills; interns and exchange students for authentic learning; integration of lang. studies with other disciplines 2. Extensive opportunities for in-country study: active exchange program; encouraging students; great attraction; sizeable proportion take it often on scholarship; working closely with exchange partners and Study Abroad office to avoid student's discouragement; biggest exchange program in Univ. for Asia (Appendix); significant contribution in alignment with Univ. policy (International Educational/Strategic Plan); students prefer shorter exchange programs: development of intensive exchange program with a univ.: effective, flexibility 3. Collaborations with university Japan club: student-run; head of department mentor; strong connection with club; special club: dynamic intercultural (lang. students, exchange students, Japanese students); extremely active; engaged in important events; recipient of awards multiple times
--	--

4. Course advice: attendance to course advice sessions **to promote** Jap. And other langs.; not effective due to **low student attendance**; still strong presence at International Studies course advice sessions; offer advice through different means: facebook, email, interview, etc.

5. Focused recruitment of Chinese students: noticeable presence of Chinese students in classes; research done of other Australian Univs.: general trend; **flyer to attract students effective**

6. Promotion of Diploma in Languages: discontinued program but important since it allowed anyone to take lang. study; a year before discontinuation, “we finally succeeded in including the “add language to your degree” icon in Univ. website

- **10 External recruitment strategies**

Strategies to raise public awareness about offering and students enjoyment:

1. Open day: always active in important events; e.g. Japan Club; raising awareness about “vibrant and internationalised student culture”

2. JALTA: understanding school sector; one member of staff joining the Japan Association for language teachers of Australia and vice-president for 2 yrs: most useful for **connections with high schools:** important stakeholders in unit offerings

3. Speech contest: long-term participants + host for Japan Foundations’ Speech contest: **attracts students and teachers** from secondary/tertiary education institutions

4. Facebook: power of social media; **promotion tool;** increasing visibility; connecting alumni w/ current students

5. Harmony day march fundraising: students and graduates collecting donations supporting tsunami victims; reported **in press;** video made by Japanese and Media Arts student in youtube

6. Fundraising: bigger fundraising event bilingual poetry reading, night food market, movie; uniting community (univ. and beyond); **money raised; advertised in press;** Head of Dep.’s article in free-press for Japanese-Australians to promote fundraising events hosted by program

7. Nara Festival: popular festival; encourage students to participate; number of people; students volunteer; **promotion** of program in community

8. Engagement with High School: principal’s requested help resulted in collaborative relationship; 4 forms of **collaborations** so far:

1. student assistant supervised weekly by HoD; worked very well; work integrated learning

2. Show and tell session about Jap. Crafts to **attract attention to potential students** about Jap. And industrial design area in Harmony day

3. Promotion event organized by graduate working as Jap. Teacher at school with Japan Club involvement

4. sharing Jap. teaching resources to enhance teaching and enhance attractiveness and usefulness of their library; donation of materials

-Further collaborations in discussion: promotion of their excellence (Appendix); through Univ. key programs and connecting students

9. Collaboration with AFDA: Australian Defence Force; small number of students from there; graduates teach there; **growing demand** due to important military ties between Australia and Japan; e.g. NDP of Japan and AFDA about to sign an exchange program; good Univ. position for **formal promotion of program** to AFDA; 15 yrs of accommodating students; students speak highly of program; students

end up teaching there; in conversation formalisation of Jap. Study at Univ. as part of exchange program (NDP + AFDA): **financial benefit** but most of all for Univ. public profile

10. Collaboration with Design area: program in Faculty of Arts and Design; Japan known for excellence in design and aesthetics; not offered at ANU; to **attract high quality students and for benefit of the Faculty:** some focus in curriculum; new exchange relationship with a Jap. Design Univ. in preparation

• **“Classes spread over several days”**

-emerged in previous review; recommendation implemented with **no success:** classroom and staffing availability; survey conducted; reinstating of traditional weekly format + special arrangements for “non-attenders”

• **“Teaching costs vs income”**

-Past efforts in the improvement of financial viability: 4 strategies:

1. Increasing revenue

2. Reducing costs: reduction of class hours; increased teaching responsibility of continuing staff; management of casual staff; contributing to other areas

3. Other financial contributors

4. Research income

5. Consultancies

6. Scholarships, fellowships and travel programs for Univ. students

7. Exchange partnership and contribution to UCELI’s operation

8. Japanese exchange partners bring income to Univ.’s ELICOS

WAYS TO STRENGTHEN PROGRAM:

• **Increasing enrolments**

-Maintaining existing strategies: **keep increasing necessary;** have increased during last 3 yrs; upward trend = good strategies

-New Initiatives (4) for student recruitment:

1. Communicating our forte (= against competitor):

-less presence than ANU; different structures and academic focus: ANU focus on acquisition of lang. proficiency; more than double contact hours; program perceived as less serious; but program’s advantage: **attract students with less time**

-Program’s goal: complementing, enhancing students’ degrees; not to produce language experts but develop Asian literacy to function in multicultural world; must be **communicated to prospective students, partners and course convenors**

2. Collaboration with Univ. Secondary Schools:

-long history of collaboration; teaching assistants and graduate as a teacher; strengthen ties in future + create new model of lang. study curriculum; arrangements:

1. UC, UC School, FUE collaboration: teachers from Japan coming to the school as part of the exchange program with partner Univ. in Japan; inspire more interest in Japanese lang. learning

2. Joint events with UC Japan Club and UC Schools: host events: Jap. Animation movie evenings + Jap. Embassy = enhance presence; most importantly: current students meet prospective students = **powerful promotion strategy**

3. New model of streamlined language learning: more fluid learning arrangement, e.g. gifted students accelerating learning via attending higher level classes; easily achievable; minimum adjustments

3. Targeted marketing –international students

1. **Attracting Chinese international students:** has fuelled rapid growth in Jap. Classes around Australia; in first yr. units as elective; potential to grow in this area

2. Jap. Language as a recruiting feature for the Asian market

-**Univ. looking to increase Chinese students in future;** closure of program = popular elective no longer available; less attraction for students to come to Univ.: in program not as severely affected by limited English skills as in other areas

-Jap. Pop culture extremely popular in China; strong affinity; Jap. + Engl. = more job opportunities; increases attractiveness to study at UC

-Additional potential targeted market: Korean: largest number of Jap. Students per capita (Australia, second); populate higher level units

4. **Design focus:** close relation to design: some students' specialization, situated in the Fac. Of Arts and Design

-**Advantage in relation to ANU:** studying design in combination with Jap.; results from students survey on motivation for taking up Jap. point potential of Jap. and design to grow together. Action:

-Kobe Design University collaboration

-In the works with staff from Dep. Of Design. Delegation at Univ. and invitation to talk to President on collaboration projects; four potential areas: research, staff exchange, application for UNESCO city of design, undergraduate exchange

-Undergraduate exchange: **enhance recruitment for Jap. and Faculty** (taking up subjects other than Jap.); exchange students could be taken to schools

• **Improve retention rates**

-willingness to increase retention after first yr and to accept students willing to study early units

-6 new strategies:

1. **Addressing unit level issues**

-many students from other Facs. wishing to enrol in Jap. do it in their second yr = **unable to pursue to levels after Jap. First yr.** due to 30cp limit; **more flexible** alternative: reintroduction of Special Studies to allow them to continue

2. **Student participation rates**

declines gradually; students' overestimation for self-learning + cramming; self-reviewers recommend regular low value assessments in and out of class time

-immigration regulat. For internat. students: **75% attendance may seem authoritarian** but need to be reminded to help avoid dropping out
-making class time effective, interesting and efficient: strong teaching + blended learning stratgs.

3. **Development of student's language learning skills**

-Different from other subjects; skills need to be learned; **Research** from ANU: three areas to help retention:

-First strategy: Teaching about language learning and what to expect; done in TESOL program; potentially adapt them to **online lectures** for new students in all langs./core subject in lang. degree if reintroduced

4. (Second strategy:) **Improving mentoring program**

	<p>-Provide mentors for students who are falling behind: already in place; could be formalised for weaker students thus monitored more systematically</p> <p>-Senior Chinese students of Jap. mentor first yr. ones to help understand English and Australian academic culture (evidence of difficulty in questions about assessments via emails)</p> <p>5. (Third strategy:) Course content adjustments</p> <p>-Reduction of hours and course content done in alignment with univ. budgetary needs and higher demands on (working) students</p> <p>-E.g. adjustment of contents in Jap. 1A: complete katakana in second semester</p> <p>-E.g. online delivery of small revision quizzes; potentially incorporation in existing online engagement</p> <p>6. Accommodating needs of differing levels of students</p> <p>-mix of true beginners with false beginners problematic: raised in research by ANU study; current trialling of group work for oral practice for mutual benefit; outcomes to be analysed at end of yr.</p> <p>• Reducing costs</p> <p>-Strategies to reduce face-to-face hours (most obvious way to reduce costs) without reducing quantum of permanent staff</p> <p>1. Restructuring face-to-face hours</p> <p>-Replacing face-to-face hours with online engagement</p> <p>-Simple replacement with online does not work well, part. early stages; students not engaged well, part. when not part of summative assessment (not graded)</p> <p>-If pursued, more attrition</p> <p>-If reduction necessary, done in control: when expected low physical attendance; when time pressures, online engagement; weeks when this happens, hours reduced, money saved</p> <p>-current class time for oral communication: done in students' own time: number of hours reduced; in-class tests: in lecture settings</p> <p>-Replacing one tutorial with a lecture</p> <p>-Replacing Jap. 1A with a lecture: total contact hours, reduce tutoring costs significantly: total hours per semester, less in second semester due to less groups.</p> <p>-Same strategy to Jap. 2, but smaller savings since only two groups.</p> <p>-Estimated annual savings in table</p> <p>2. Staff secondment to other areas</p> <p>-Helps to monetise core teaching savings; e.g. TESOL (strong connection): HoD and another staff member worked teaching TESOL subjects; HoD and another staff member could take on roles outside of teaching Jap. "should there be an opportunity for secondment"</p> <p>• Summary: Proposals for future improvement (= cost-efficient) (idem summary as in ES)</p> <p><u>ELIMINATING DEFICITS</u></p> <p>-No deficit clearly demonstrated; program committed to generate income; 4 strategies:</p> <p>1. Reintroduction of advanced units</p> <p>-Essential for 3 reasons:</p> <p>a. net income, no deficit; removal produced loss of income to Fac.</p>
--	---

b. **enhance flexibility** to program connected to **recruitment power** (Schools)
 c. likely to produce star graduates (contribution to program, Fac. And Univ. profiles)

-Financial profile of Advanced Jap. units
 -calculation of **total revenue and expense** shows in 3 yrs = income despite small enrolments (with 54% for the internal charge CAMS rate) (Common Area Maintenance?)
 -Closing caused students discontinue study at Univ. (transferred or stopped)

-Enhancing flexibility:
-Catering for different entry levels
 -to **attract wider range of students**; only students who start in their first semester can complete major; students with knowledge have to do **cross-institutional**; less attractive option (time-consuming; timetable difficulties)
 -Proposed major structure (pre-existent before closing down units)

-Language education in ACT schools
-Research: entry level significant factor for **future recruitment** for program and Univ.
 -State Schools offer to teach 8 langs.; Jap. most popular, most widely taught; due to Fed. Gov. program, Jap. students increased 253% in three yrs. = **increase demand** to continue at Univ. level
 -Benchmark of total Jap. learners against other langs.: Jap. increasing
 -Benchmark by year level: in Jap. larger cohort in younger years = greater demand for Jap. study and partic. for advanced courses
 -Summary of secondary schools teaching Jap.: not compulsory in yr. 11-12; amongst the ones that offer, all but one offer Jap.
-Evident future demand: crucial to have major structure for **financial health** of program and for Univ. student recruitment
 -Anecdotal evidence: school leavers like combination of Jap. study with other subjects; **closing down Jap. program = losing students**

-Producing star graduates
 -advanced units produced alumni with a variety of achievements: 8 examples
 -2 reasons: extensive major option = **attracts serious students; success warranted**; complements specialization of student
 -authentic materials: examples; current issues; engagement in a major project combining skills from Jap. and degree; 6 examples, including creative media arts pieces + journalism piece

2. New offering

1. New non-academic courses

-Another pathway to **generate revenue**; potential demand after closure of certificate in languages in another local Univ.; careful planning since **closure due to low enrolments**: identify **potential clients and target** them

2. Formal collaboration with ADFA

-**potential client** which has been for 15 yrs. Due to **program flexibility; more demand now** due to national strategic needs; exchange program with Japan's defence academy; ongoing negotiation for making Jap. study at Univ. part of exchange program
 -alternative option due to timetable difficulties: intensive course for them; HoD in discussion with Major at Academy

3. Teacher retraining intensive course

-ANU offers during normal semester schedule

-**Increasing demand** of Jap. at school level = demand of qualified teachers: **potential demand** for intensive retraining since easier commitment for them

4. Offering a graduate certificate course

-Advanced units cancelled (despite co-taught; no extra costs other than 2 hr. per sem. for convenors); summary of **income generated** in 3 yrs: small **enrolments**; **effective earner**; students left; **lost income** “entirely”

-Closure of local univ.’s certificate + gov.’s focus on Asia: **potential demand** to grow; previous grad. diploma 4yrs.; difficulty to commit due to fluidity of workforce in society; proposal of graduate certificate, 2 yrs.; potentially to be extended (as TESOL’s model) to a diploma/linked to International Studies Masters

-Japanese for Business in Winter Term

-potential interdisciplinary venture with another Fac.; fits into alternative degree structures with elective in WT

-goal of program; coverage; focus of the language; option to continue mainstream Jap. course

-for students: **Intern. Students** with alternative degree structures/ students wanting to enhance career opportunities

3. UC as the base for Cool Japan in Australia

-Japan’s government cultural export initiative to increase Japan’s presence globally; **budget**; in Oceania no projects yet; recommend Univ. to be the base in Australia and Oceania; benefits: increase **recruitment power locally and internationally**

-3 examples possible collaborations

-HoD in conversation with Embassy for this; very positive response from Minister of Economy and embassy; potential extremely exciting and attractive collaboration

4. Reviewing financial modelling for internal charges

-**language teaching more costly** than other subjects (cannot be taught at a lecture theatre); fed. gov. funds; summary of funding loading

-Summary demonstrates student load should not be taken on face value: 1.5 times per student; lang. lab, equipment at other Univ., not this one

-Lang. teaching more viable when adjusting financial model: reflecting gov. support, extra 50% funding vs base-funding cluster which covers the extra cost; internal charge should discount this extra money.

-Calculation presented: much fairer, more true to Gov. intentions internal funding model; urge Fac. To change current model into this one

Summary: Increasing profitability (idem summary as in ES)

• **Concluding Statements**

a. Program’s contributions

-Program’s success in aligning program with both the Univ.’s as well as the Federal government’s strategic goals

-A1: 1. International Education Policy: text taken from the Univ. policies (online address directed to a restricted access page)

A3. Breakdown of continuing and discontinuing students –from Jap. 1A to Jap. 1B

A4. Summary of USS data for Jap. units (2009-2011) benchmarked against university rates and showing higher rates than Univ.’s average

	<p>A5. Revision of generic skills learning outcome A6. Unit content adjustment (2009-2011) A7. Univ.'s Asian region exchange programs- Jap. dominance A8. Chinese international students in Japanese language courses in Australia A10. Arts and Design faculty- unit student numbers A11. Degree studied by Jap. language students</p>
<p>Electives (range of offerings) associated with short-term financial gains</p>	<p>-financial viability increases with recruitment of single elective students</p>
<p style="text-align: center;">THEMES FROM INTERVIEW DATA Main themes where most of these would fit in:</p> <p>1. Commissioning of review = instrument to disguise authoritarian decision-making (because of fear to public exposure about authoritarian decision-making) 2. Predetermined decision to close program associated with macro efficiency management (getting rid of languages and letting another university absorb students; university broader goals) overrules micro-efficiency</p>	
<p>Focus on efficiency</p>	<ul style="list-style-type: none"> • Efficiency: Fixation with small enrolments A206. [...] Data really didn't get through anyway. I had this in front of him and pointing out he said 'oh, but it's not big enough number' and just kind of— it was extremely frustrating and infuriating in every sense. • Efficiency widespread practice: University principle guiding other strategic university operations, e.g. hiring managers who use same principle. Example: past review four years earlier: Q85. Did she er commission it? A86. Yeah, she brought in some external people. Er I think she already had agenda so, you know, when dean was brought in probably already had a mission of basically save money. So saving money might mean flexible delivery and reduced face to face hours um reduction of (xxx) numbers, offerings and cut some staff, and make the budget better. I think there was a— Q87. The main drive [for— A88. Yeah,] absolutely. It's all financial. • Managers' fixation with efficiency leads to ignoring students/university priorities A350. Students, students' experience was the least of their concern. A352. Oh, well, that's what I wanna know. That's kind of I got— gone to the complete despair because they say like um they talk about, you know, internationalizing students as a strategy plan and they talked even about exchange and we were doing exactly that and they just kind of think 'oh, yeah, that's important but not as important' and just kind of take strategy, vision, policy and I used that as a guiding line: Teaching content I— there are several things I believe students should learn but as a guiding principle of the organization, the program within the organization, I was using that as a guiding principle and they just said 'eh, okay', so I mean, what can we do? -Fixation on efficiency's blinding effect:

	(A): “Um well, it’s a perception of ‘we are losing money’” which in (A)’s view was true in the past for some languages but unfounded in the current time
Commissioners	<ul style="list-style-type: none"> • Personal factor A204. [...] they don’t take a, well, I shouldn’t say ‘they’ but the particular dean I dealt with, the particular senior management I had at that time, [...] they just don’t face up the data. • Top-down practice = Servitude/uninformed decision-making/authoritarian A206. Er this particular dean, the new dean, his role is do what vice chancellor tells him to do, regardless of data. • Restricting review process K64. And then, also writing to the faculty. Now this is a closed review. We weren’t allowed to submit to the review directly er all we were allowed to do was to writ- to write a letter of concern to the dean and to the vice chancellor, and then, indirectly provide assistance with their response to the review. [...] but the university made it ver- the faculty made it very difficult to establish the true financial position, and then, in terms of the criteria of the review, was very strict er I should, I mean, it was very unfair, it was very strict in, in the criteria it’d established, the parameters it’d established of the review. It was— did it in such a way that it was impossible to actually enter a lot of information. It was very, very restrictive er which is, you know, I’ve— for me, it was very unfair. K.68. [...] but he [the dean] carefully managed the review in such a tight way that nothing was made public and that er in the end, you know, there were restrictions placed [...]
Biased review process	<ul style="list-style-type: none"> • Trigger (unprofessional, unethical, dishonest) -Delay predetermined authoritarian decision to close program down Q229. So then after that you guys got active. A230. Active, and so they said ‘oh, okay, we will think about for— this semester we’ll open (L6) and we’ll think about it’ and 20 of March, the terms of reference is this. -Commissioner’s face saving, embarrassment (K): K68. Well, it was the dean in this particular case, and the reasons why this dean did it this way— he was forced to do it as a result of public outcry because it became— the decision, the decision was made to close the language without discussion, then, there was public outcry in the press and he was forced to establish a review er to satisfy public concern and then— [...] — it was saved but with certain face saving er criteria for the dean. -Commissioning of review = Simulacrum A264. So the whole— my position was whole review was bollocks. • Biased ToR

	<p>A6. [...] and the (L6) language program review was pretty outrageous: The review document um what was it, what's called? The term of reference was all heavily critical, it wasn't neutral, it was saying something like um (L6) program was, well, we had 5 reviews in the past and in all of them (L6) program was told to make various changes and (L6) program has been refusing to make any of those changes and then not in— and so forth.</p> <p>Q7. Was that true?</p> <p>A8. No, it wasn't and [...]</p> <ul style="list-style-type: none"> • Biased purpose: review used as pretext not only to deal with finances but specifically to push a false argument, that is, a <i>deficit</i> in finances: <p>K64. [...] And one of the issues, of course, was the financial aspect of it. The driver for the review was, you know— it was supposedly costing money, losing money. Turned out, of course, it wasn't losing money</p> <p>K62: [...] There ((It)) was a review conducted by the faculty, with the aim of potentially closing a language program</p> <ul style="list-style-type: none"> • Biased questions <p>A242. Money]. All— it's all attacking, it's not question. It's kind of—</p>
<p>Deficient management practices</p>	<p>-Important distinction in attrition (electives/degree) and profound causes of attrition not acknowledged</p> <p>A178. They didn't look] very carefully but they see the first year number and the second year number and 'look it's so different' and they don't see why it's different, they don't see how many students left the [target language] studies because of the course structure and all of that. They don't see that kind of thing.</p> <ul style="list-style-type: none"> • Dean's decision to close down program before review: -Rushed and unexpected; surprising for program staff and detrimental for students; unfair <p>A2. [...] The (U4) initially proposed to close (L6) program, out of blue er in 2— in December 2011. He said 'oh, we will finish this year so we're not gonna offer next year' in December. It was already December 14 or something. Er no, it's not even that, it was December— it was 20 or something. It was few days before Christmas.</p> <p>A6. And the students were already enrolling in. Students er choosing (U4), in some cases, because they can study something and (L6), and suddenly being told and there was no precursor whatsoever.</p> <p>(C): unfair and cowardly top-down decision to closing down program: no consultation to program staff</p> <ul style="list-style-type: none"> • Sloppy management <p>A8. No, it wasn't and there wasn't 5 reviews even. They, they don't even keeping track of number of reviews. [That's ridiculous.¹²⁰</p> <p>-----</p> <p>A248. [...] and then first time they send me the document, they forgot to add terms of reference and that kind of sloppy, sloppy thing and—</p> <ul style="list-style-type: none"> • Commissioner's denial of factual information

A204. Well, they never thought of it. I pointed out and they just kind of ‘oh, yeah, but you have still small numbers’. **They just don’t take any information**, they don’t take a, well, I shouldn’t say ‘they’ but the particular dean I dealt with, the particular senior management I had at that time, **they just don’t face up the data**.

A328. [...] but **I resent the fact that they refused to acknowledge the information, they refused to acknowledge the facts**. They didn’t take—

Q329. The evidence.

A330. They just really— **most sobering thing is they don’t care about evidence**, they don’t care about fact. That’s— that was really, **really awful**. ((laugh))

• **Corruptive practice: commissioner’s self-interested selection of reviewers:**

A12. Er it’s] someone from faculty. It’s, it’s supposed to be from er **one internal faculty** back then dean— **associate dean of education** and er in different de- **dean of different faculty and another person is er bringing from the external**. She is from— someone from (C3) actually um **was a mates of dean**. So it was kind of— it’s er **committee made to make his decision happen**.

• **Covering up a lie: Fabrication of data and delaying access to study teaching costs**

- (A) contributes a lengthy but illuminative account highlighting
- a) **difficulties to access** data to the extent of having to resort to threatening in order to get it, or making a legitimate demand such as requesting data look like a backdoor affair;
 - b) **lying** involved in the assertion made in the Review Notice in view of the absence of data;
 - c) **rushed**, last minute assembling of data

A270. Uhuh. Okay, for instance, the important thing: Money, **teaching course income issue. I have never been given the information** and er when **I was told (L6) program has been losing a lot of money for many years, I was very surprised because no one said that**. (L6) was breaking even, just we have to be careful (xxx).

A274. [...] Anyhow, so I’ll show you the money: I went to— so **when I was told that, I went straight away to the business manager of the faculty asking for financial information** of the (L6) program. She said ‘**oh, we don’t have such things**. We have only the degree and the faculty level but we don’t have the (L6) program, or languages’ and then **what the hell the dean is talking about? He was talking from nothing**.

Q279. Badly managed.

A280. And also **lying bastard**. How can you tell (L6) program is—

Q281. Losing money.

A282. ‘**Losing a lot of money**’, they say. So that’s how bad it is and I—

Q283. So lies, they were lying.

A284. They are lying, yeah, basically, lying. Well, what's **created lying, so I demanded financial information** and they said 'oh, it's, it's not (xxx), it isn't sufficient, and it's—' I said 'no, you said that ongoing deficit and a large amount of deficit. **I need information**' and **they refused to provide and they gave me the information, financial information only few days before the deadline of this response.**

Q285. Are you saying that they did have information?

A286. **Actually, they didn't have it but so they have to make it but they didn't want to give me because they were wrong.**

Q287. After you requested it, they said 'we don't have it' but after you requested it, they thought we better get something together, put something together here.

A288. Well, they didn't, so **I went straight up to the registrar** of the university and **'this is outrageous, and if they claim that and if you conduct a review based on the financial problem without giving me financial information, how on earth am I supposed to— they will really compromise university's credibility. (xxx) I'm gonna go on the media'**. **And then he— the registrar acted on that.** Er I got information just 2 days before, but I—

A290. But, well, I kept all my record, so I knew who I hire as a tutor, how many hours, and then, I can look up the EB, the Enterprise Bargaining documents, so I can see how much the tutors cost was, and I know how much on cost is, and I know how much I earn, and I know what level my colleague is, and half time, so I calculated all from first principle so I— by the time they gave me information **I had all analysis done and ready to roll but they didn't give me any information, I had to go backway, and for the revenue I had to go to the university statistics office to get data because—**

A292. Uh, no, that statistics department understood what's happening and they said that that's outrageous so **they quietly gave me the information.**

Q293. Because the financial office refused.

A294. Yeah, and the financial office through faculty gave me the information 2 days **before the deadline.**

These findings also supported by (C) who believes that

C7. [...] In the case of er of (U4) er for example, [...] we discovered quite quickly that the figures on which they were basing er their er their decisions at the level of the dean, the **figures were inaccurate** er very, very considerably inaccurate.

C75. [...] and they had **bad figures**

• **Assumed low professional standards in university management**

(A) relishes the idea of the commissioner's surprise to receiving an unexpected long and in-depth self-review report:

Q247. Did you address [all of these questions?]

A248. Yeah,] probably. **Well, I wrote 84 pages document for that ((laugh)) which they didn't expect. They gave me 2 pages [...]** so this is executive summary.

	<p>-Lack of trust in /respect to management’s professional standards Research work in self-review began before review was commissioned out of suspicion: meeting with dean announcing closure of program</p> <p>A340. Well, when that happened, I immediately start writing about the (L6) program um before this review happens. I wrote— I was— Q341. Before which review you wrote it, 2011? A342. After 2011 they said— 2011 they gave me a decision. Q343. Yes, they gave you the decision, that’s right and then during the first semester, they decided to conduct a review, yeah, correct. A344. Yeah, but in between I thought ‘I should write up how wonderful (L6) program do because these idiots probably doesn’t know’ [...]</p>
<p>Negative emotions</p>	<ul style="list-style-type: none"> • Lack of respect to management A8. No, it wasn’t and there wasn’t 5 reviews even. They, they don’t even keeping track of number of reviews. [That’s ridiculous. • Frustration - A206. [...] it was extremely frustrating and infuriating in every sense. -A330. They just really— most sobering thing is they don’t care about evidence, they don’t care about fact. That’s— that was really, really awful. ((laugh)) -A352. Oh, well, that’s what I wanna know. That’s kind of I got—gonne to the complete despair [...] • Bitterness A328. [...] but I resent the fact that they refused to acknowledge the information [...] • Vengeful -(A) relishes the idea of the commissioner’s surprise to receiving an unexpected long and in-depth self-review report: Q247. Did you address [all of these questions? A248. Yeah,] probably. Well, I wrote 84 pages document for that ((laugh)) which they didn’t expect. They gave me 2 pages [...] so this is executive summary. A274. Yeah, also um they thought yeah, that’s one. 2 women: One is Australian, one (L6) and Australian one is only half time. She doesn’t have a research behind her to defend her and I’m (L6) so they might have thought we were rolled over easily. Too bad. ((laugh)) A422. [...] but I was the hardest one to deal with in all language staff. I know how to fight. • From trusting management to Loss of faith in university A344. [...] and then I still had faith in the system. I thought ‘he’s new, he’s out of his depth, he doesn’t know, so if I explain [...]
<p>Confidence</p>	<ul style="list-style-type: none"> • In program A344. Yeah, but in between I thought ‘I should write up how wonderful (L6) program do [...] so if I explain what contribution (L6) program is making— financially I didn’t have the information but er like what students did, the um— were coming those (D3) (L11) language course students, and er contribution to the er exchange

	<p>program. (L6) is the second biggest destination in the (U4) as an exchange destination, and second to (COU5), right? And er (CO1) area exchange was actually nonexistent without (L6) program, it was 90% or 96% was (L6) program in (AL) area.</p> <p>• In meeting expectations / compliance A490. [...] So what I learnt was I ticked all of the box— Q491. Yeah, that would be my last question. A492. I ticked all the box. All pdf, I was never being questioned ‘yeah, you’re doing great, great, great’ and the student number increased 3 times from 2004 to 2011. It was increasing.</p> <p>A494. And our teaching evaluation was really high like usually 90%, 100% student satisfaction rate in every unit. That’s how well we run it. A496. Um so I’m very confident it wasn’t the problem of the program. It was a very useful thing, financially not losing anything, even bringing in money and er (L6) language most widely taught language in (S/T1), and all the information, all (xxx) information I told them in manifold, and a lot of graduates write to the council and then saying how they thought— how useful that was to their lives, they probably received 40 letters.</p>
Reviews perceived as threatening instrument	A2. Okay er while I was at the (U4) which was from 2002 to 2013, we had quite a few reviews. Language was always under pressure. Um when I joined the (U4) I already sensed all the language people are feeling that university is gonna come and get them, so review is used for that.
CASE STUDY 7 (Non-Go8) Year: 2010	
<p style="text-align: center;">THEMES FROM REPORT</p> <p>-Aligning program with institutional/governmental regulatory priorities with main priority proving strategies are being put in place and planned via promoting what exists and making necessary changes to attract more students using a variety of commercial strategies (marketing/promotion/ monitoring attrition) to keep and increase enrolments</p> <p>-Responding to student needs</p>	
Willingness to comply	<p>-Program reflects priority of univ. education: equip students for global context; languages = assets for access to knowledge, expertise and leadership; = investments in engagement with global markets and economies (cited references)</p> <p>-pattern:</p> <p>-in program coherence: emphasis on program takes institutional framework seriously and orients program towards meeting goals established by framework. E.g. in “Content, workload and assessment”: program’s design “supports the development and integrations of research experience, and skills, experiential learning and academic literacy” according to the “Teaching and Learning Framework”.</p> <p>- very first sentence in each subarea within academic design starts with assertion ensuring program meets expectations: “The graduate qualities are well integrated”; “The teaching/research nexus is a primary focus of the [program name] program”; “From 2008, the applied linguistic major began to introduce on-line courses.”; and “Strong focus on indigenous content in key courses”</p>
Alignment with priorities	<p style="text-align: center;"><u>PURPOSE OF THE PROGRAM; STRATEGIC CONSIDERATIONS</u></p> <p>• Current and future contribution of the program to University, Division, and School strategic goals:</p> <p>1. “University’s strategic emphases”: outstanding learning outcomes, quality and richness of student experience, students’ adaptability and contribution as global citizens; strong nexus between teaching and research</p>

2. **“University’s current goals 2010-2012”**: international engagement; teaching and learning quality

3. **“the Corporate Plan 2010-2012”**: incorporation of experiential learning; provision of Personal Learning Environments

4. **“the Division Plan 2010-2012 and the School’s priorities”**: teaching and learning quality, particularly to student satisfaction, incorporation of experiential learning and scholarship of teaching and learning (= research component)

Program’s contributions to Priorities

-International and intercultural education experience

-Collaborative inquiry

-Mutual respect and understanding

-Content of teaching informed by research (Research Centre for Languages and Cultures: members teach and coordinate program; with strong research backgrounds on language and intercultural communication; national leader; examples of projects undertaken)

• **Extent to which the program meets stated aims and objectives**

-Description of program:

Overall aim and objectives against Graduate Quality traits: aims defined according to attributes

Program Structure: majoring and sub-majoring, range of languages (some available with cross-institutional arrangm.), applied linguistics and languages integrated; specializations; foci of program: applied and research; possible inclusion of overseas study; number of enrolments in lang. majors (= broad range of programs enrolling)

--conclusion: overall meeting of aims; needs improvement: specializations, particularly, Translating and TESOL

• **Relationship to other programs in university**

-Current and potential connections: e.g. ESL/ Master of Teaching (external stakeholders offered to promote it)

• **Emergence of any new markets and how the program may exploit those markets**

-Domestic and international

-Targeting students (marketing plans designed to attract students)

2.1. Demand and market position

(Main topic: marketing as main action to increase enrolments)

2.1.1. Market analysis

-5 tables

-Benchmarking with competitors: 2 local univs.

-Benchmarks: CEQ; total language enrolments; TER per degree in 3 univs. where langs. taught; languages offered

2.1.2. Language offerings

2.1.3. TER

2.1.4. Promotion

2.1.5. Other programs in (name of state)

2.1.6. Other programs in Australia

2.1.7. Features of similar programs

2.1.8. Issues

2.1.9. Other threats

-**Discussion**: emphasis on promotion of program through marketing

-Background: external demands: regulated environment (government regulation?)

-Description of marketing activities at other univs. used to support promotion

	<p>-Comparison between program and other programs at local univs.: deepening understanding of program but focus on enrolments</p> <p>-Issues and other threats focus on strategies that help attract students: eg. Increase range of langs.; increase electives in AL; strategies to attract international students; increase visibility of program “to grow student enrolments”; monitor state regulation affecting detrimentally enrolments</p> <p>2.2. Proposed Actions</p> <p>- 6 actions: PA1, PA2, PA3 “marketing”; PA4 “lack of visibility” = marketing; PA5, PA6 attracting students: expanding range of particular langs. with potential markets in partnership with other univs.</p>
	<p style="text-align: center;"><u>PROGRAM VIABILITY</u></p> <p>3.1. Student Profile, Equity and Inclusivity</p> <p>-1 table (international student enrolments in program across 5 yrs.)</p> <p>3.1.1 The Current Student Profile of the Program</p> <p>3.1.2 Access and participation of equity groups</p> <p>3.1.3 Any other Data/ Information Relevant to the Standing of the Program (e.g. stakeholder views, age of program, strategic importance, relation to competitor programs)</p> <p>3.1.4 Any trends that may indicate improvement</p> <p>Discussion: some on student participation (= %) of student profiles and backgrounds but main orientation clearly enrolments; e.g. in 3.1.1. ends with existing marketing plan seeking stronger connection with schools to increment enrolments (domestic and international local students + students from international markets)</p> <p>3.2 Proposed Actions</p> <p>-4 actions: focus on enrolments</p> <p>-PA4: continue to implement domestic and international marketing plan with domestic strategies outlined in PA1 (linking program with career paths (external stakeholders support it) for promotion purposes); PA2 (explore new specializations, e.g. Translating/Interpreting and TESOL) & PA3 (language major for Australian minority group)</p> <p>-Timetabling problems for students wishing to study cross-institutionally not addressed in PA nor low enrolments (but in final PA section it is said it needs investigating).</p> <p>-Language: participation = enrolments</p> <p>-Conclusion: Viability associated with enrolments</p>
	<p style="text-align: center;"><u>RETENTION AND SUCCESS</u></p> <p>-Brief</p> <p>-Discussion: gaining control of student numbers and aligning with priorities (surpassing divisional retention rate)</p> <p>-positive (retention rate increased)</p> <p>-negative (only exceeded division rate in 1/3 yr. = have to exceed division rate?; a concern given positive student feedback); need to surpass divisional retention rate; possible causes from focus groups: feeling lost within range of options</p> <p>4.1. Proposed Actions</p> <p>Focus: avoiding student attrition</p> <p>-3 actions: PA3 (strategies to assist students at risk); PA1 (strategies for monitoring retention and success: inform students about options); PA2 (further research: causes for attrition/ students failing courses)</p>
	<p>5. Current Student Satisfaction</p> <p>-3 tables</p> <p>-Discussion: detailed explanation of student satisfaction results from various student evaluation instruments showing overall student satisfaction except in two courses</p>

	<p>5.1 Proposed Actions</p> <p>-1 action: PA1 program's commitment to comply with institutional priority: continue to monitor closely and maintain upward trend in alignment with [...]; identify causes for low rating in two courses through program reviews (= PROGRAM REVIEWS FOR IDENTIFIED PROBLEM: PART OF BUREAUCRATIC SYSTEM ASSOCIATED WITH PROBLEMS)</p>
	<p style="text-align: center;"><u>GRADUATE SATISFACTION AND OUTCOMES</u></p> <p>-4 tables 6.1 Graduate Satisfaction and GCEQ performance 6.2 Graduate Destinations and Employment</p> <p>-Discussion:</p> <p>-benchmarks CEQs against national averages; match between CES & CEI very high rates; CEQs scores compared to scores in 4 other universities which conform one group of universities; CEQs compared to other groups of Australian universities; student comments = all justifying high level of satisfaction</p> <p>-survey on graduate destinations and employment: benchmark full time employment against national average; benchmark educating professionals against divisional average; employment pathways unclear to students; actions suggest informing students; helping students networking with industry; integrating employment skills and information management in program.</p> <p>6.3. Proposed Actions</p> <p>-5 actions</p> <p>-Emphasis: MARKETING: better 'sell' program to maintain and increase student satisfaction thus ensuring student attraction to program:</p> <p>-Matching program with particular career pathways/possibility of postgraduate studies: 5 PAs PA2, PA3, PA4: activities in conjunction with state organizations to match program with employability PA1, PA5: program activities: follow up on alumni/ integration of informational and employment skills component into program ("Career Management")</p>
	<p style="text-align: center;"><u>EXTERNAL STAKEHOLDER SATISFACTION</u></p> <p>7.1 Identify major external stakeholders: 7.2 Professional Accreditation 7.3 Formal articulation and credit arrangements</p> <p>-List of names of external stakeholder groups -Maintaining relationships: existing relationships and developing new ones for the purpose of involving stakeholders in connecting graduates with job market - 9 Key points from discussions with groups: marketing and promotion: -3,4,5,9: marketing and promotion of program -2: promotion of Master of Teaching ensuring students' continuation of language study (previously emerging in strategic considerations and viability) -1,6,7: connecting program with career options: 1 (guiding to plan choices and career connections); 6/7: involvement of representatives from two employment organizations to attract the contemporary student (same actions as in viability: promotional orientation to attract students) -8: emphasis on 1 graduate's perspective: value of program but highlighting how well program prepares students beyond its geographical location -Professional accreditation empty, no explanation -Names of partner universities and processes for reviewing and maintaining arrangements and relationships -4 providers for cross-institutional arrangements: 3 from Asia and one local (specialization: business and technology to capture international students studying at that institution); 1 potential TAFE under consideration (Interpreting and Translating) -maintenance + reviewing processes under review; program and school involved; actions below</p>

	<p>7.4 Proposed Actions</p> <ul style="list-style-type: none"> -Underlying subject in actions = enrolments demonstrating how teaching team contributes to program sustainability -PA 1-5 ensuing from stakeholder consultation: <ul style="list-style-type: none"> PA1: ensure students in M. of Teaching may continue language study PA 2, 3, 5 organizing/improving promotion activities PA4: career info sessions to ensure fit bet'n program and employment -PA 6-7: promoting program locally and internationally
	<p style="text-align: center;"><u>ACADEMIC DESIGN</u></p> <p>8.1. Program Coherence</p> <ul style="list-style-type: none"> 8.1.1 Content, Workload and Assessment 8.1.2 Teaching and Learning Arrangements 8.1.3 Program Objectives and Graduate Qualities Profile 8.1.4 Issues <ul style="list-style-type: none"> -pattern: emphasis on program takes institutional framework seriously and orients program towards meeting the goals established by framework. E.g. in “Content, workload and assessment”: program’s design “supports the development and integrations of research experience, and skills, experiential learning and academic literacy” according to the “Teaching and Learning Framework”. -Focus of 8.1.3 on correlation between the two; program “highly congruent with the Graduate Qualities”; how program tackles these; examples in many courses, above all in AL -2 issues: 1) Lack of data from other awards; 2) Gaps: lack of content and overlapping of assessment and focus; suggested actions <p>8.2. Academic Design</p> <ul style="list-style-type: none"> 8.2.1 Graduate Qualities 8.2.3 Teaching/Research Nexus 8.2.4 Flexible Environment 8.2.5 Indigenous Content <ul style="list-style-type: none"> - Same patterns as above: very first sentence in each subarea starts with assertion ensuring program meets expectations: “The graduate qualities are well integrated”; “The teaching/research nexus is a primary focus of the [program name] program”; “From 2008, the applied linguistic major began to introduce on-line courses.”; and “Strong focus on indigenous content in key courses” - Teaching/research nexus discussion focus: AL; for languages: timid statement asserting this aspect “is also an element of other parts of the degree”; examples only for AL -Goal: make delivery available totally on-line for AL; for langs.: mixed-mode but emphasis on flexible delivery opportunities offered by the newly acquired “state of the art” “teaching and learning facility” = asset for program and used also for marketing; timeframe for total integration -Stress: integration of indigenous content in various courses all connected with applied linguistics; collaboration with a specialized indigenous centre within the univ. (media article: trend: dissipate/remove specialized centres and blend them into faculties; this centre disestablished after review identified pitfalls: declining enrolments first one) <p>8.3 Program Support</p> <ul style="list-style-type: none"> -Emphasis: “active” approach by the program “in liasing” with two University resources: the library; univ. unit to support students to complete courses + counselling (highly valued in text); table: number of contacts in 5 yrs. showing increase (weakening sign of educational standards? Not raised) <p>8.4 Proposed Actions</p>

	<p>- 5 actions: 2 (flexible delivery); 1 (expand range, partic. Indigenous); 2 (program support: strategies in line with Univ. program for supporting students + work with library with acquiring resources for new facility)</p>
	<p style="text-align: center;"><u>SUMMARY OF ACTIONS/CHANGES PROPOSED</u></p> <p>-Summary of PAs and implementation: main focus = attracting students/retaining students</p> <p>-12 recommendations divided into four sections:</p> <p>1. Participation (4 Rs) (assoc. with student numbers and profiles):</p> <p>-R1 (11 actions; R with most actions): Enrolment numbers + first preference enrolments and TER cut off: goal = recruitment of domestic/international students</p> <p>-R2, 3, 4 (1 action each): extending pathways and specializations</p> <p>2. Employment prospects (6): focus on ensuring relationship between program and career options</p> <p>-R5 (4 acts.): focus on relationship with external stakeholders and promotion</p> <p>-R6 (1 act.), R7 (1 act.), R8 (1 act.), R9 (1 act.), R10 (1 act.): focus on ensuring students are ‘marketable’</p> <p>3. Retention and success (1 action):</p> <p>-R11: investigation of threats: low participation in lang. courses; action connected to actions in R1 = monitoring retention to avoid enrolments drop/decline</p> <p>4. Courses (2 actions):</p> <p>R12: goal: improve core educational content; review of two specific courses; emphasis on addressing data from CEI/SETs from those courses (student satisfaction = part of “University Strategic emphases” (student satisfaction associated with enrolments)</p>
	<p>10. Sign Off</p> <p>-“Three layers” or “points of accountability” (confirming interview data)</p>
	<p style="text-align: center;"><u>APPENDICES</u></p> <p>-3 appendices</p> <p>1. Marketing Plan: Bachelor of Arts (Languages and Intercultural communication), MBAL, 2011-2012</p> <p>-Designed in collaboration with “Marketing and Business Development Management International” = focus on marketing</p> <ul style="list-style-type: none"> • Objectives: increasing enrolments and TER for the Bachelor of Arts • Measurables: enrolments numerical data accord. to student profile; goal = increase by 30% type grown in preferences to date; projected target growth for international students (why/implications) <p>-Domestic</p> <p>-International</p> <ul style="list-style-type: none"> • Situational analysis (context setting): current situation for attracting students -Domestic Market: increased market in school leavers due to past efforts; further actions for marketing -International Market: perfect storm points (negative impact in international enrolments); potential to increase in collaboration with state education body; FPOS rates for past seven yrs. • Market opportunities: key opportunities to market the program -Domestic market: 4 actions; e.g. professional development of teachers (so school teachers can promote program at their schools) -International market: attract FPOS from five specific countries from biggest market to new markets; two recommendations: making program attractive to International students (e.g. linking to specific postgraduate pathways, etc.); employ 1.6 full time staff dedicated to Local International Market • Key messages: topics when communicating about program and actions to make program competitive

	<p>• Marketing activities; A year in review-Domestic Marketing Activity -Summary of past marketing activities: date, target audience, marketing activity, outcomes, comments</p> <p>• Future Planning: International and Domestic Marketing Activity for MBAL 2011-2012 -Summary plan: date, target audience, marketing activity (reflecting key messages), objectives, comments; ongoing character of plan: two timing blocks + evaluation and review of marketing program</p> <p>2. Supplementary data to evidence program support: letter by external to the program stakeholder prove level of collaboration (activities and strategies) between program and Univ. centre for student support</p> <p>3. Supplementary data to evidence program support: Program connection with Library: letter by external to the program stakeholder with information about library services to different campuses; library holdings, academic and research support, etc. Expenditures for Langs. and Linguistics holdings of past 4 yrs.; expenditures specifically for program</p>
Alignment purpose of program reviews	Compliance with university guidelines: statistical alignment as a result of review, two 'program reviews' recommended focused on issues at a more granular level: "students' feedback as part of the review of the curriculum, content and teaching". However, discussion states goal is for teachers to "ensure that their own SETs are not below Divisional average" thus revealing a compliance purpose for program reviews as well as for accreditation reviews.
Pressures	Examples: -surpass divisional retention rates -maintain upward trend of student satisfaction
Issues for students	-timetabling for cross-listed study (not addressed) -feeling lost within range of options -increase of use of counselling and student support to complete courses (not addressed)
THEMES FROM INTERVIEW DATA	
Evaluation purpose = internal accreditation: compliance with internal regulation to get internally 'accredited' (continue/not continue?)	<p>R8. [...] in my university, every three to five years, a program must be evaluated formally by its teaching team, so, all faculty contribute. That particular evaluation occurs within the structure of a school. [...] Now, every year, the university itself, and my university is one which is possibly more regulated than others, okay? It is because of the former vice chancellor that we had who was very, very precise on these matters. Every year the university gathers data on the program. It is numerical data only, but it looks er at things like total enrolments in the program er retention er those kinds of indicators, okay? So we have a picture, a snapshot every year. If the, if it seems to be normal shall we say, there's no problem, it just continues. Every five years there must be, regardless of whether the program data are okay or not, there must be a review-</p> <p>Q9. A full]-</p> <p>R10. -A full evaluation. There is a policy, our policy is (RP1). It has a number and a name and it is to do with it-it occurs in every department. It's not just for the languages section. This is a standard university policy and procedure, and every five years we have to do that.</p>
Participatory self-evaluation approach	<p>(R) confirms that "The teaching team, all of those involved, are part of the evaluation"</p> <ul style="list-style-type: none"> - collaboration goes beyond data collection, and data analysis and involves the writing of report which in turn solves the issue of authorship: <p>R12 [...] And we wrote up the sections of the report. The drafts were all discussed in committee, everybody participating.</p>

	<p>- collaboration seems to be a requirement by the institutional policy.</p> <p>R8. [...] in my university, every three to five years, a program must be evaluated formally by its teaching team, so, all faculty contribute.</p> <p>R10. [...] There is a policy, our policy is (policy reference). [...] The teaching team, all of those involved, are part of the evaluation.</p> <p>- which (R) admits to be circumvented at times but no explanation provided as to why:</p> <p>R16. [...] in some programs, the program director just sits down, looks at the data and writes it, [...]</p> <p>• (R) = organizer + participant:</p> <p>-“Um I designed that particular evaluation”</p> <p>- “And so, I designed a plan and decided which people in a group could do which parts.”</p> <p>-R10. [...] What I <u>did</u> do, my part, the part that I gave to me was to review, to conduct focus groups: focus groups with students, focus group with the staff and focus group with external, [people external]</p> <p>Q11. Stakeholders.]</p> <p>R12. Stakeholders um so that we could get their perspective on our program. How our program is viewed from the outside. Now, others did comparisons with other universities, now, whether we like it or not, we are required to benchmark, it’s called, with other universities, difficult to get data often because it is after all a competitive system, but we did that. Er others er looked at what is the situation with language teaching and learning more globally, a lot of us are involved in that, but there was a kind of literature review, a context statement, so, one person was responsible for that, someone did the benchmarking, I did the focus groups, someone else looked at just the data as it stood um and the student feedback and so on, and so different people looked at different um aspects.</p> <p>• Outcome of participatory self-review: positive: “And it worked very, very well”</p> <p>• Conditions for PSR success:</p> <p>1. Leadership:</p> <p>- trusted and with social skills “to find a way” to deal with “disruptive forces”, that is, with “others who say ‘look, I don’t want to touch this’” which (R) in fact accomplished “we certainly did in that process, I can say.”</p> <p>- someone for whom ‘collaboration’ is a value of the work culture, which in fact, expressions such as “<i>We</i> are monitoring this trend closely [...]”; “<i>We</i> have developed a series of [...]”; “<i>We</i> are also negotiating ways of articulating [...]”; “In the domestic market, <i>we</i> have recently targeted [...]”; and “<i>We</i> are also planning a pathway through [...]”, are an indicator of a well-established collaborative culture at that program and therefore the “disruptive forces” may have been perhaps the only few that took convincing.</p> <p>- ‘collaboration’ had to be worked on; support to administrative leader officially responsible for carrying out review and taking over to lead the work:</p> <p>R34. Well, that’s where,] yeah, no, it’s not remunerated. That’s where I worked with the program director to be absolutely honest, the program director was very new and inexperienced, and he said ‘(P5), I don’t know how to begin’ and I</p>
--	---

said 'well, never mind. **What we need to do is make sure that everybody contributes**, I think that we should do a plan' and I sat down and I wrote it with him. I actually wrote the plan. I said 'now, what you need to do is approach each one, and so he went and said 'are you prepared to do this section exactly as I had written? I made a judgment about who could do what, and so, I used the different talents in the best way and then of course when he introduced, he chaired the group, when he chaired it, he's inexperienced, he said 'what if they all say no', and I said '**well, I will be there and I will say that in my view we need to]**' and so, that's how it happened. And so it's dialogue, it's dialogue, but I pushed the dialogue from within.

- with key notions of evaluation in general and participatory evaluation in particular such as the importance to involve internal members of staff with experience in reviews, knowledgeable about benefits specifically of participatory evaluation and with an understanding that staff are not going to get any reward for doing such work, and therefore that commitment to review needs to be built:

Q35. Did you mention the **advantages** of doing such work?

R36. Well, yes, yes, **I certainly, I certainly put the argument that** collaboration was the way to do it, that we needed to take collective responsibility, that this was not the responsibility of any one individual, that the improvement of the program needed to be the improvement of the program as a whole, not just one or two courses. So, I told the story of those kinds of benefits, surely. **Now, what would have happened if I had not intervened** with my less experienced colleague? I don't know. **There would have been more of a struggle, I'm sure.** But then, you know, **people like me exist in different institutions as well. So, you know, you use the collective force of who's there.**

- highlighting rationale for selecting an evaluation participatory model and the importance of sharing this knowledge with staff in order to encourage and create commitment and other key benefits

"this must be collaborative, contributory, we have to have everyone working on this' which is further extended in the following explanation

R16. [...] I wanted it to be, you know, **I proposed to my colleagues that it** needed to be much more participatory, everyone needed to be involved and contribute, that **we needed to own and take responsibility and that we needed to take responsibility for the program, in fact, and the changes that we-because** unless people have been part of the dialogue they may or may not be committed to those recommendations, and so, you know, in order to, and this is why I think these kind of processes need to be participatory, **because that is how you gain buy-in**, as we say in English, in other words, commitment to the outcomes of the evaluation. **If not, they would say** 'okay, so why should I bother? This has nothing to do with me. Nothing to do with me, I'm not going to bother'. **And so you build ownership and responsibility and also through the dialogue** mm?, through the dialogue we begin to put out onto the table the different points of view, the different experiences, **because**, you know, I with a longer history of experience with these things will see it in one way, someone else who has been with us for two years will see it with totally different eyes, and **we need to capture those different perspectives.**

- awareness of distinction between internal and external approaches to evaluation. E.g. insights about (a) the importance of an evaluator's awareness of the implications of an external approach; (b) the benefits and better suitability of an internal participatory model for evaluating language programs in contrast with an external approach:

R6. [...] I mean, you can go there as an outsider but **you must recognize that you are an outsider to the context** and people have huge investments in their own program, people do not just work in a program, they invest their whole intellectual energy and their physical energy into the program. It is based on their own belief about what a language program is and how it works and so on, and when it is an entirely external affair, there is a risk that the people on the inside really feel disenfranchised, that they are concerned that maybe people from the outside will not understand properly what their issues really are, or that we don't understand the history strongly enough, or that we perhaps don't care about history because evaluation looks back but really it is with a view to looking forward, and so, **people find that quite difficult** and I think that I would find it **confronting** too. Um my own view on evaluation is that **it really does need to be um participatory, collaborative** if we are going to get the best results that **people from the outside can't understand the program as people who are living that program will understand it**. And that intelligence needs to be brought there.

- and (c) the contrast between evaluation "for compliance" versus "meaningful" evaluation highlighting how compliance can be turned into an opportunity for improvement:

R110. [...] In the example that I've given from my own program, we could have treated that exercise as an exercise in pure compliance: 'we are required to do this and we shall. We tick. We have done it.' If it is just an exercise in compliance we've lost the meaningfulness that we were talking about. Not worth doing.

- awareness about educative function of evaluation ensues from participatory evaluation, e.g. exchanging information during meetings with external stakeholders:

R30. [...] so **we educate them as part of involving them** in the evaluation. Evaluation can have an **educative function** at many levels.

- awareness of importance of adding an external element to an internal evaluation in order to avoid bias:

R8. [...] And then, there is the evaluation um in my own university, which is an evaluation which is um conducted totally internally, well, largely internally but there is an external er element.

But what does this "external element" involve? According to (R), various committees of internal members of the university check and eventually endorse report prepared by program (evidence in sign off section of report):

R8. [...] So, the teaching team conducts the evaluation, but where does the accountability come in? The report of the evaluation is submitted to the head of school, it is considered by the school board. After it is considered by the school board, it may be returned to the teaching team for amendments or further work, that may happen. After that, when it is endorsed by the school board, it goes to the division board, okay? So, the dean of the division, here it would be-at (U2) it would be faculty. So, the school board, faculty board. We don't have faculties, we have divisions, okay? So school board, faculty board. Once it leaves the faculty board, again the faculty may request further revisions or further explanations of the evaluation. You can't change the facts of the case, you know, but you can change er the judgments that you make about it or they may ask for further explanations. And then it goes to the um academic um board of the university as a whole, okay? So there are three layers, there are points of accountability.

	<p>However, can these internal “points of accountability” be considered “external”? The following quote discussing the ‘public’ or ‘private’ nature of the evaluation report shows that for (R) they are:</p> <p>R12. [...] That report that you have seen is not a public document, but nor is it private and internal because of the three accountability check points: the school, the division or faculty and university wide. So, the eyes of other members of the university community who had nothing to do with our languages program, were going to be looking.</p> <p>Thus, are these “accountability points” what make this report not private? And why are reports not public? Isn’t this university being run mainly with public funding?</p>
Diversity of institutional culture	<p>R16. [...] And you will see that what comes out in there is the institutional flavour because the flavour of the (U8) is going to be different from the flavour of the (U2), it’s different from the flavour of the (U11), because each institution has its own culture, if you like, and so, our different institutions will require different styles of er evaluation but that is mandatory, we must do that.</p>
High stakes review policy	<ul style="list-style-type: none"> • Justification for focus on marketing, promotion enrolments, attracting students <p>R14. Absolutely, well, they would have a say as to whether this program was effective, viable, viability mm? Now, for us it’s a small program and so the vulnerability for us is, it’s not on quality, the results show very high on quality but it’s boutique, it hasn’t been there for a long time, it doesn’t have a long history, and so, there’s a strong section on marketing, we need to market the program more, and so, well, you’ve read the report, [...]</p> <ul style="list-style-type: none"> • (R) describes the institutional policy driving the review, and states if annual “snapshot” of program “seems to be normal shall we say, there’s no problem, it just continues”. But what makes the numerical data appear as “normal”? Does this mean that if the numerical data are not “normal” the program doesn’t continue?