

Senior secondary school-based assessment: quality management processes  
and teachers' professional learning

Brent Atherton

M. Ed. (Research and Innovation) (Flinders), B.Sc. (Adelaide), Dip. Ed.  
(Adelaide), Dip. T. (Adelaide Teachers College),  
Grad. Cert. Sc. Ed (Chemistry) (Adelaide)

Thesis submitted for the degree of Doctor of Philosophy

School of Education

Faculty of Education, Humanities, Law and Theology

Flinders University

2015

## Table of Contents

Table of Contents .....	2
List of Figures .....	9
List of Tables.....	10
Abstract .....	12
Declaration .....	14
Acknowledgements .....	15
Chapter 1: Introduction .....	16
1.1 Increased use of SBA.....	16
1.2 Quality management of SBA .....	17
1.3 Quality management terminology.....	17
1.4 Quality management processes .....	18
1.5 Learning from quality management processes.....	18
1.6 Quality management processes in the SACE.....	19
1.7 Research purpose .....	19
1.8 The researcher's position .....	21
1.9 Research goals .....	21
1.10 Chapter summary and thesis outline .....	22
Chapter 2: The Research Context and Literature Review.....	24
2.1 Introduction.....	24
2.2 Possible teacher learning.....	26
2.3 The context of the research .....	32
2.3.1 The international context.....	32
2.3.2 The Australian context .....	39
2.3.3 The South Australian scene.....	42
2.4 Empirical evidence for teacher learning from quality management processes .....	48
2.4.1 Peer-reviewed claims linking social moderation and teacher learning ...	49
2.4.2 Other evidence about teacher learning associated with social moderation.....	51
2.4.3 Recent evidence for teacher learning from social moderation.....	54
2.4.4 A summary and analysis of the evidence for teacher learning from quality management processes .....	62
2.5 Quality management of school-based assessment .....	64
2.5.1 Quality assurance of school-based assessment .....	64

2.5.2	Quality review of school-based assessment .....	66
2.6	Effective professional learning for teachers.....	66
2.6.1	Terminology associated with teacher learning.....	67
2.6.2	Effective professional development activities.....	68
2.6.3	Effective teacher learning.....	72
2.6.4	Adult learning.....	80
2.6.5	The quality of teacher learning.....	84
2.6.6	Teacher characteristics associated with effective teacher learning .....	88
2.7	Summary of research questions .....	92
Chapter 3:	Methodology and methods .....	96
3.1	Research purpose .....	96
3.2	Methodology .....	97
3.3	Teacher invitation .....	103
3.4	Research methods and sources.....	103
3.4.1	SACE Board documents.....	104
3.4.2	Observation of quality management processes .....	104
3.4.3	Questionnaires .....	106
3.4.4	Teacher interviews .....	125
3.4.5	SACE Board officers focus groups .....	134
3.4.6	Data collection overview.....	135
3.5	Delimitations.....	136
3.6	Limitations .....	137
3.7	Ethical matters .....	140
3.8	Evolution of the research .....	140
3.9	Preliminary treatment of the data and approaches to data analysis .....	141
3.9.1	SACE Board documents.....	142
3.9.2	Observations of quality management processes.....	143
3.9.3	Treatment of missing data from questionnaires .....	143
3.9.4	Coding of interviews and focus groups.....	144
3.9.5	Evidence of teacher change.....	151
3.10	Summary.....	155
Chapter 4:	Results: Teacher involvement .....	156
4.1	Participants.....	156
4.1.1	Teachers .....	156
4.1.2	SACE Board officers.....	159

4.2	Teacher involvement in the quality management processes .....	159
4.2.1	Teacher involvement in planning forums.....	160
4.2.2	Teacher involvement in learning and assessment plan (LAP) approval and clarifying support panels .....	161
4.2.3	Teacher involvement in clarifying forums .....	161
4.2.4	Teacher involvement in moderation.....	161
4.2.5	Increased teacher involvement? .....	161
4.3	Reasons for teacher involvement in the quality management processes ...	164
4.3.1	Reasons for involvement in planning forums .....	164
4.3.2	Reasons for involvement in LAP approval and clarifying support panels.....	164
4.3.3	Reasons for attending clarifying forums .....	165
4.3.4	Reasons for involvement in moderation.....	166
4.3.5	Reasons for involvement in the quality management processes .....	167
4.4	Teachers' school characteristics and involvement in the quality management processes .....	167
4.5	Professional and gender teacher characteristics and involvement in the quality management processes.....	168
4.6	Communities of practice and the quality management processes.....	168
4.6.1	Participation in the Physics online forum .....	169
4.6.2	Intra-school collaboration .....	170
4.6.3	Inter-school collaboration .....	171
4.6.4	Encouragement of communities of practice in the quality management processes.....	171
4.7	Use of exemplars.....	172
4.8	Overview of teacher involvement.....	173
Chapter 5: Results: Form of the quality management processes .....		174
5.1	The quality management processes as effective professional development activities .....	174
5.1.1	Active learning in the quality management processes .....	175
5.1.2	Coherence in the quality management processes.....	177
5.1.3	Duration of the quality management processes.....	178
5.1.4	Collective participation in the quality management processes .....	179
5.1.5	Conclusions about quality management processes as effective professional development activities .....	180
5.2	Features of effective teacher learning in the quality management processes .....	180

5.2.1	Learning opportunities embedded in teachers' work .....	180
5.2.2	Moderation opportunities during the year .....	181
5.2.3	Access to expertise .....	183
5.2.4	Challenging prevailing discourses .....	185
5.2.5	Overview of the quality management processes as effective teacher learning .....	188
5.3	Features of effective adult learning in the quality management processes	188
5.3.1	Quality management processes: teacher input into planning .....	189
5.3.2	Quality management processes: relevance to teachers' needs .....	189
5.3.3	Congruence of the quality management processes with teachers' prior experiences .....	191
5.3.4	Use of appropriate materials and activities in the quality management processes.....	193
5.3.5	Conclusions about the quality management processes as effective adult learning.....	197
5.4	Opportunities for deep learning in the quality management processes .....	197
5.4.1	Active learning .....	198
5.4.2	Building on prior knowledge.....	198
5.4.3	Confronting teacher misconceptions .....	199
5.4.4	Discussion with other teachers .....	199
5.4.5	Discussion with presenters .....	200
5.4.6	Overview of opportunities for deep learning in the quality management processes.....	201
5.5	An overview of the form of the quality management processes.....	201
Chapter 6: Results: Content of the quality management processes .....		203
6.1	Subject focus of the quality management processes .....	203
6.2	Awareness of the reasons for change .....	203
6.3	Performance standards: purposes, advantages and disadvantages.....	204
6.4	Learning opportunities in task design .....	205
6.4.1	Planning forums .....	205
6.4.2	Clarifying support panel.....	206
6.4.3	Clarifying forums .....	206
6.4.4	Moderation panels .....	207
6.4.5	Exemplars.....	207
6.4.6	Feedback to teachers .....	209
6.4.7	Communities of practice .....	212
6.4.8	Summary of learning opportunities in task design.....	214

6.5	Learning opportunities in grading.....	214
6.5.1	Planning forum.....	215
6.5.2	Clarifying support panel.....	215
6.5.3	Clarifying forum.....	215
6.5.4	Moderation panel.....	217
6.5.5	Feedback to teachers .....	218
6.5.6	Communities of practice .....	219
6.5.7	Exemplars.....	220
6.5.8	Summary of learning opportunities in grading .....	222
6.6	Education re bias.....	223
6.7	Overview of the content of the quality management processes.....	223
Chapter 7:	Results: Teacher learning .....	225
7.1	Teachers' descriptions of learnings from the quality management processes .....	225
7.1.1	Planning forums .....	225
7.1.2	Supervisors .....	228
7.1.3	Clarifying forums .....	229
7.1.4	Moderation panels .....	231
7.1.5	Feedback after moderation .....	235
7.1.6	Exemplars.....	237
7.1.7	Communities of practice .....	238
7.1.8	A summary of participants' perceptions of learning from the quality management processes .....	241
7.2	Evidence of teacher change.....	241
7.2.1	Teacher awareness of the reasons for change .....	242
7.2.2	Changes in teachers' assessment beliefs .....	243
7.2.3	Familiarity with SACE Board terminology .....	249
7.2.4	Consistency in application of performance standards.....	254
7.2.5	Teacher assessment confidence.....	255
7.2.6	Teacher attitudes .....	257
7.2.7	Overview of evidence of participant change.....	258
7.3	The nature of teacher learning .....	258
7.3.1	Brenda's learning journey .....	259
7.3.2	John's learning journey .....	265
7.3.3	Kevin's learning journey.....	270
7.3.4	Participants' reflections on learning.....	275

7.3.5 Overview of teacher learning .....	280
7.4 Teacher learning and teacher characteristics .....	281
7.5 A summary of the evidence for teacher learning .....	281
Chapter 8: Discussion .....	283
8.1 Purpose of research .....	283
8.2 Number of teachers involved in social moderation .....	284
8.2.1 Comparing the SACE and Queensland systems .....	285
8.2.2 Why not the Queensland system? .....	287
8.2.3 Moderation in a high-stakes environment .....	288
8.3 Duration of teacher involvement in social moderation .....	294
8.4 Teacher learning from the quality management cycle .....	294
8.4.1 Intra-school collaboration .....	295
8.4.2 Extra-school collaboration .....	296
8.4.3 Feedback .....	300
8.4.4 Exemplars .....	300
8.4.5 Ongoing teacher learning .....	301
8.5 Teacher learning from induction programs .....	302
8.5.1 Providing a rationale for change .....	302
8.5.2 Challenging teacher preconceptions .....	303
8.5.3 Teacher participation in induction planning .....	304
8.5.4 Duration of an implementation program .....	304
8.5.5 Learning and mandated change .....	306
8.6 First Year Evaluation of the new SACE .....	307
8.6.1 Performance standards .....	307
8.6.2 Exemplars .....	308
8.6.3 Assessment task design .....	308
8.6.4 Moderation feedback .....	308
8.6.5 Clarifying forums .....	309
8.6.6 Other topics mentioned in the evaluation and by participants .....	310
8.6.7 Summary of the first year evaluation .....	311
8.7 Summary and recommendations .....	312
Induction programs .....	312
Quality management cycle .....	313
Social moderation .....	313

Chapter 9: Conclusion.....	315
9.1 Key Findings .....	316
9.2 Contribution to the literature.....	318
9.3 Limitations of the study .....	319
9.4 Suggestions for further research .....	321
9.5 Reflections on the research process .....	322
9.6 Concluding Statement .....	323
Appendices .....	325
Appendix A: Teacher invitation .....	325
Appendix B: Officer invitation.....	326
Appendix C: Teacher letter of introduction .....	327
Appendix D: Officer letter of introduction.....	329
Appendix E: Research information sheet.....	330
Appendix F: Teacher consent form.....	332
Appendix G: Officer consent form .....	333
Appendix H: Questionnaire 1 protocol.....	334
Appendix I: Questionnaire 2 protocol.....	339
Appendix J: Questionnaire 3 protocol.....	340
Appendix K: Interview 1 protocol.....	344
Appendix L: Interview 2 protocol.....	348
Appendix M: Interview 3 protocol .....	351
Appendix N: Verbalisation protocol.....	356
Appendix O: Own task protocol .....	357
Appendix P: Round 1 focus group protocol .....	358
Appendix Q: Round 3 focus group protocol .....	362
Appendix R: Questionnaire data analysis .....	365
Confidence in formative .....	365
Confidence in summative .....	365
Formative practice .....	366
Summative practice .....	367
School support .....	368
Appendix S: Final coding for teacher learning .....	371
References .....	377



## List of Figures

Figure 1: Relationship between concepts addressed in literature review .....	26
Figure 2: Traditional model: professional development through changing teacher knowledge.....	27
Figure 3: Guskey’s model: professional development through changing teacher practice.....	28
Figure 4: Desimone’s model for the effects of professional development .....	28
Figure 5: Clarke and Hollingsworth’s (2002) interconnected model of teacher learning .....	29
Figure 6: Experiential learning cycle based on Kolb (1984) .....	30
Figure 7: Traditional and Guskey learning models incorporated in the interconnected model (Clarke and Hollingsworth, 2002) .....	30
Figure 8: Dual scale format used in James and Pedder’s teacher questionnaire.....	110
Figure 9: Adam's learning regarding LAPs explained using the interconnected model of teacher learning .....	153
Figure 10: Change sequence Brenda—consulting other teachers.....	260
Figure 11: Change sequence—enactment of externally modelled practice .....	261
Figure 12: Change sequence Brenda—exemplar use in task design.....	262
Figure 13: Change sequence Brenda—exemplar use in task design.....	263
Figure 14: Change sequence John—the importance of task design.....	266
Figure 15: Change sequence John—student information from performance standards and using performance standards to write reports .....	267
Figure 16: Change sequence Kevin—assessment of practical work Phase 1 .....	271
Figure 17: Change sequence Kevin—assessment of practical work Phase 2 .....	272
Figure 18: Change sequence Kevin—belief in usefulness of LAPs .....	273
Figure 19: Clarke and Hollingsworth’s interconnected model of teacher learning .	306

## List of Tables

Table 1: Features of effective teacher learning .....	73
Table 2: Knowles's assumptions about adult learners .....	81
Table 3: Knowles's process model for adult learning .....	82
Table 4: Designing teacher professional development using adult learning principles (Loucks-Horsley et al., 2003) .....	83
Table 5: Methods used for each of the quality management processes .....	101
Table 6: Schedule of SACE Board quality management processes .....	102
Table 7: Items related to teacher confidence in the formative use of SBA.....	111
Table 8: Items related to teachers' beliefs about the purposes of moderation .....	115
Table 9: Items related to confidence of teachers in their own assessment practices	116
Table 10: Statements related to teachers' practices in the formative use of SBA....	117
Table 11: Statements related to teachers' summative assessment practices.....	118
Table 12: Items related to school expectations and support, from Rosenholtz (1989).....	120
Table 13: Items related to school support for SACE Board involvement.....	121
Table 14: Items used to identify transformative learning phases.....	123
Table 15: Data collected in each round of questionnaires .....	124
Table 16: Questionnaire dimensions and research questions addressed.....	125
Table 17: Overview of interview topics .....	129
Table 18: Data to be collected in each round of interviews .....	134
Table 19: Relationship between data and research questions .....	136
Table 20: Alternative strategies for consolidating codes .....	149
Table 21: Coding organisation based on research questions and sub-questions.....	150
Table 22: Percentages of teachers in educational sectors .....	157
Table 23: Overview of quality management processes as effective professional development activities .....	174
Table 24: The quality management processes as effective adult learning .....	189
Table 25: Opportunities for deep learning in the quality management processes ...	198
Table 26: SACE website materials 2011 and 2014: assessment design advice .....	208
Table 27: Section of moderation feedback sheet .....	210
Table 28: Frequency of grades submitted by teachers online prior to clarifying forums .....	216

Table 29: SACE website materials 2011 and 2014: grading of student work .....	221
Table 30: Performance standards for AE1 .....	221
Table 31: Participants' opinions on helpfulness of the planning forums (n = 9).....	226
Table 32: Physics teacher assessment confidence (percentages) from planning forum exit surveys (n = 101) .....	227
Table 33: Participant responses to the clarifying forums (n = 13).....	230
Table 34: Moderators' opinions of the educative benefit of moderation (n = 6).....	232
Table 35: Participants' meanings of <i>criteria</i> in interviews (n = 12).....	250
Table 36: Examples of specific features and performance standards .....	252
Table 37: Participants' use of the term <i>performance standard</i> (n = 12) .....	253
Table 38: Grades for verbalisation exercise.....	255
Table 39: Frequency of steps in learning .....	276
Table 40: Frequency of external stimuli in change sequences.....	279
Table 41: Items related to teachers' confidence in the formative use of SBA .....	365
Table 42: Statements related to teachers' practices in the formative use of SBA ....	367
Table 43: Statements related to teachers' summative assessment practices.....	367
Table 44: School support items from Rosenholtz (1989) .....	369
Table 45: Items relating to school support for SACE involvement .....	370
Table 46: Reliability of school support data .....	370
Table 47: Coding for Research Question 4 .....	371

## ABSTRACT

2011 was the first year in which the South Australian Certificate of Education (SACE) employed social moderation of school-based assessment in all Year 12 subjects. Social moderation had been recommended in the 2006 report of the Ministerial Review of the SACE. The panel recommended greater reliance on the professional judgements of classroom teachers and had been impressed by Queensland's system which, while verifying standards, served as professional development for teachers. Much literature on social moderation makes similar claims for its educative value. A major purpose of the research reported in this thesis was to examine evidence for such claims. The *new* SACE<sup>1</sup> envisaged social moderation of school-based assessment as one process in a quality management cycle. The present research was, therefore, an exploration of the potential for teacher learning from involvement in a range of processes, including teacher meetings, membership of assessment panels, moderation feedback and the use of assessment exemplars. The research focussed on teachers of Physics, a subject with no history of social moderation in the SACE. It was considered that these teachers might be more likely to evince learning because it was their first year of involvement in an assessment system that included social moderation. The data included SACE documents, observations, a series of questionnaires and interviews with thirteen Physics teachers over a twelve-month period and two focus groups with SACE Board officers. Teachers reflected on their revised assessment tasks and the means of their development; they also graded student work using "think aloud" protocols. The data were examined in the light of the literature on effective teacher learning and adult learning.

It was found that the quality management processes possessed some of the features associated with effective professional development and teacher learning. Increased teacher involvement in quality management processes, viewed as desirable by the Review Panel, appeared, however to have been minimal. Teachers reported changes

---

<sup>1</sup> The revised SACE was initially known as the *future* SACE (*future* SACE Office, 2008), which was changed to the *new* SACE (SACE Board of South Australia, 2010n; SACE Principals Partnership Strategy Group, 2009) as implementation approached, with both terms italicised. The non-italicised term "new SACE" will be used subsequently.

in their assessment practices, particularly in tasks requiring students to undertake practical work or research an issue with social or environmental implications. The findings suggested that the changed assessment practices arose from changed syllabus requirements rather than from the ongoing cycle of quality management processes *per se*. Teachers reported that changed practices commonly involved experimentation followed by reflection on, and evaluation of, the tasks. Through these processes the teachers seemed to demonstrate experiential learning, reflective practice and deep learning. Membership of a community of practice contributed to constructive reflection, particularly where there was more than one Physics teacher in a school.

It is argued that, while a quality management cycle, such as used in South Australia, has the potential for teacher learning, learning would be more likely if the processes are specifically designed to facilitate such learning, and teacher involvement is encouraged. It was concluded that, for learning to be more likely, the processes should be based on evidence in the literature about (a) effective professional development and adult learning strategies and (b) research where social moderation has been found to contribute to teacher learning.

## **DECLARATION**

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Brent Atherton

## ACKNOWLEDGEMENTS

This thesis is dedicated to Mum and Dad and Irena.

Mum and Dad encouraged me from the beginning of my learning journey. Irena encouraged me in the final stage but contributed much more than encouragement—listening, asking questions, discussing, transcribing, proof-reading—and did all so willingly.

Alan and Kerry, my supervisors, were essential. They encouraged, but indicated what needed to be done; they always found another reading to help complete the picture. They persevered with the task of attempting to turn a science teacher into an academic writer.

There were many who assisted at various stages over the last few years. Brian, Cliff, Graham, Julie, Koula, Neville, Rob and Ron come to mind. Thanks to the Chemistry teachers—Alison, Luci, Philip, Sally, Susan and Winsley—who helped by trialling the online questionnaire.

Anne carried out many transcriptions and did much proof-reading, both of which were deeply appreciated, while Trevor brought the eye of an English teacher to corrections of the later versions. Any errors that remain, however, are mine.

I am thankful that the SACE Board provided sources that I had not envisaged at the beginning of the research and am especially appreciative of the officers who were willing to put aside some of their busy time to offer their reflections in focus groups.

Finally, I am grateful to the fourteen teachers who willingly gave their time. Without them, the research would never have happened. It was clear that concern for their students lay behind their becoming teachers, their seeking to improve their teaching, and was the reason they volunteered for this research.

Brent Atherton

# CHAPTER 1: INTRODUCTION

The current research explored the potential for teacher learning in quality management processes associated with school-based assessment (SBA).<sup>2</sup> The research differed from much previous research in two ways: it examined a range of quality management processes rather than only social moderation and it was undertaken in the context of high-stakes senior secondary assessment rather than the earlier years of schooling. The research examined the quality management processes in the South Australian Certificate of Education (SACE) in 2011 when a new system of assessment, common to all subjects, was introduced. The new system required SBA as the major component, the adoption of performance standards and social moderation of the SBA component. These changes provided the opportunity for the research conducted for this thesis.

This chapter outlines the international, national and local contexts. It then defines key terms pertinent to the thesis and describes the changes made and why these provided an opportunity for research into potential teacher learning in the quality management processes. After a statement of the researcher's position, the chapter concludes with the research goals and the thesis outline.

## 1.1 Increased use of SBA

Recent decades have seen increased use of teacher judgements in senior secondary assessment in many educational systems (Briseid & Caillods, 2004; Davison, 2004; Luyten & Dolkarb, 2010; Matters, 2006; Mercurio, 2008; Wolf, 1995). Teacher judgement in SBA entails planning assessment, designing assessment tasks and marking student work. In senior secondary assessment, most countries now use a combination of SBA and an external examination (Briseid & Caillods, 2004; Mercurio, 2008). All Australian education systems incorporate SBA in senior

---

<sup>2</sup> The term *school-based assessment* is used in this thesis because it was used in the SACE Review (Crafter, Crook, & Reid, 2006) and, initially, by the SACE Board of South Australia (subsequently referred to as the SACE Board) (SACE Board of South Australia, 2010f). It is used in some other states (ACT Board of Senior Secondary Studies, 2012; QSA, 2010e; VCAA, 2013). Recently, the SACE Board has used the term *school assessment* (e.g., SACE Board of South Australia, 2014a).



secondary assessment (Matters, 2006). While South Australia had incorporated SBA in end-of-secondary assessment for decades (Mercurio, 2003), changes were introduced in SBA in response to the SACE Review (Crafter et al., 2006).

## 1.2 Quality management of SBA

Although there has been an increase in the use of teacher judgements, SBA is not held in the same regard as external examinations which are more ‘trusted’ (Maxwell, 2006, p. 3). Consequently, incorporation of teacher judgement into high-stakes assessment, such as the SACE, has included quality management processes to “establish and maintain public credibility” (Mercurio, 2008, p. 2).

## 1.3 Quality management terminology

Processes aimed at ensuring the quality of assessment may operate at different points in the system. Australia’s National Quality Council (NQC) oversees quality assurance in vocational and technical education and is charged with ensuring national consistency (National Quality Council, 2006). The NQC uses the term *quality management* for the overview of the processes (National Quality Council, 2008). Maxwell (2006) pointed out that there is no common terminology in use for the different processes, suggesting that “one way of distinguishing different components is to say that *quality assurance* occurs before, *quality control* occurs during, and *quality review* occurs after the assessments are completed and the results reported” (Maxwell, 2006, p. 3, italics added). The NQC uses the same three terms (National Quality Council, 2008, p. 4). *Quality assurance* establishes “appropriate circumstances for assessment to take place” (p. 41). *Quality control* refers to monitoring assessment results and “making adjustments” (p. 41) when necessary, while *quality review* involves retrospection, “to make improvements for future use” (p. 41). While *quality review* is uncommon, some (e.g., Gipps, 1994a; Harlen, 1994; Wiliam, 1992) use the terms *quality assurance* and *quality control* in the same way as Maxwell and the NQC.

The SACE Board uses *quality assurance* for the totality of processes that assure “the integrity of the assessment and reporting of student achievement in the SACE” (SACE Board of South Australia, 2010f). The NQC and Maxwell’s use of the terms *quality assurance*, *quality control*, and *quality review* as components of *quality*

*management*, however, will be adopted in this thesis because of their wider use.

## 1.4 Quality management processes

Various strategies assist teachers in quality assurance of their assessments (Harlen, 1994); for example, use of exemplars. Quality control, the confirmation or adjustment of teachers' marks commonly known as *moderation* (Brown, 1980; Harlen, 2007; Maxwell, 2006; Wiliam, 1992), may also take different forms. Statistical processes are used, as is professional judgement (Keeves, 1994; Linn, 1993). Both are used in Australian education systems (Matters, 2006). Professional judgement moderation may utilise an external moderator or other teachers. For the latter procedure, the term *social moderation* (Linn, 1993) is used in this thesis. Because various terms are used in the literature for professional judgement moderation, it is necessary to define the terms used in this thesis. Where professional judgement uses one person, the term *judgemental moderation* is used; this term is borrowed from Keeves (1994) but redefined for this thesis. *Group moderation* refers to teachers involved in discussions of *their own* students' tasks where grade adjustments may occur (i.e., quality control). Where teachers engage in discussion for the purpose of confirming or adjusting *other* teachers' results, the term *social moderation* is used. This was introduced into the SACE for all subjects in 2011. The literature also refers to teacher discussions about their judgements, without the power to adjust results (e.g., Klenowski & Adie, 2009; L. Reid, 2007; Smaill, 2013; Wyatt-Smith & Castleton, 2005). This is, therefore, a form of quality assurance rather than quality control. Although referred to as *consensus moderation* (e.g., Connolly, Klenowski, & Wyatt-Smith, 2012; Klenowski et al., 2007; Linn, 1993), in this thesis it will be referred to as *consensus negotiation*.

## 1.5 Learning from quality management processes

Advocates of social moderation commonly claim professional learning on the part of moderators. In Queensland, where it has been used for decades (Matters, 2006), most moderators saw it as "valuable professional development" (Maxwell, 2006, p. 6) and Klenowski and Wyatt-Smith (2010b) asserted "it has been shown to build teacher assessment capacity, as well as teacher confidence" (p. 115). Others see advantages of moderation discussions extending beyond the realm of assessment.

Harlen (2004a) claimed its benefits apply “to teaching and learning as well” (p. 6) while Hipkins and Robertson (2011) argued it provides “opportunities for professional learning that aims to lift student achievement” (p. 7). The SACE Review (Crafter et al., 2006) endorsed Queensland’s moderation system as “an excellent form of teacher professional development” (p. 129). Smaill (2013) contended that teacher professional learning should be viewed “as a goal rather than a by-product of moderation” (p. 1).

Quality management processes incorporating social moderation serve three goals: public credibility, reliable assessment and teacher learning. A central purpose of the research conducted for this thesis was to examine the educative benefits to teachers of participating in the range of quality management processes that include social moderation.

## **1.6 Quality management processes in the SACE**

The SACE Board has responsibility for senior secondary assessment in South Australia (*SACE Board of South Australia Act 1983*). The Board adopted a “four-phase cycle” (SACE Board of South Australia, 2010f) to assure the community of the integrity of its school-based and external assessment. Because the current research was limited to SBA, subsequently the term *quality management processes* will refer only to those applicable to SBA. While the primary purpose of the SACE cycle is to ensure the quality of the assessment, the current research was based on the premise that involvement might provide opportunities for teacher learning.

The four-phase cycle includes planning, clarifying, confirming and improving (SACE Board of South Australia, 2010i). Planning and clarifying are *quality assurance* processes, attempting to “increase dependability of teachers’ assessments ... before the assessment is made” (Harlen, 1994, p. 10). The confirming phase constitutes *quality control* as it “provides a check on judgments” (Harlen, 2007, p. 74) while improving constitutes *quality review*.

## **1.7 Research purpose**

Although social moderation is claimed to provide professional learning for moderators (e.g., Harlen, 2004a; Hipkins & Robertson, 2011; Klenowski & Wyatt-

Smith, 2010b; Maxwell, 2006), there are few reports of research into such learning (Hipkins & Robertson, 2011). Research undertaken has usually involved the middle years of schooling rather than end-of-secondary schooling. 2011 provided an opportunity to undertake research in a high-stakes environment when social moderation was adopted for all end-of-secondary SACE subjects. Teachers of subjects with no history of social moderation were seen to offer an opportunity to undertake research into claims of associated learning. This opportunity was only possible in 2011 as, subsequently, teachers of all subjects would have experience of assessment that included social moderation.

Although social moderation occurred in all SACE subjects from 2011, it was not in isolation but as one of the four phases mentioned above. Consequently, the current research explored the potential for teacher learning in the quality management processes as a whole.

While Australian states and territories have retained much independence in educational matters, succeeding Federal governments have moved to bring them closer. A national curriculum is being prepared by the Australian Curriculum, Assessment and Reporting Authority (ACARA) although “implementation and support are the responsibility of state and territory school and curriculum authorities” (ACARA, 2012). While the various jurisdictions retain assessment control, steps have been taken to compare standards across Australia in a number of subjects (ACER, 2012). A previous Federal Education Minister threatened that states would be required “to introduce public exams ... as a condition of the next schools funding agreement” (Ferrari, 2007), a threat removed with a change of government. If assessment in the states and territories is to move closer, it is suggested that moderation of SBA will be a critical topic.

It is anticipated that the research for this thesis might contribute to this discussion. First, it can add to knowledge about teacher learning in general, while addressing, in particular, the lack of evidence relating to teachers learning from involvement in the quality management processes used in high-stakes SBA. Second, the current research addresses a call for increased research on teacher assessment literacy (DeLuca & Klinger, 2010). Third, evidence of the educative value of the quality management processes might contribute to wider use of social moderation. Fourth,

the findings may suggest improvements in the quality management processes, improving their potential for teacher learning.

## **1.8 The researcher's position**

I have been involved for over thirty years in assessment with the SACE Board and its predecessors. As a teacher of Chemistry, Information Technology and Science, I was involved in group moderation whereby groups of teachers viewed samples of student work from each teacher in the group. The group reached consensus and marks were confirmed or adjusted as agreed. When the Chemistry syllabus was reviewed in 1998, teachers were asked their opinions of group moderation. While teachers endorsed its effectiveness as professional development, there was concern over teachers being involved in moderation of their own assessments, adding a “degree of competitiveness in moderation meetings” (McGaw, Warry, & McBryde, 1975, p. 32), with some teachers assuming the role of advocate for their students. Teachers described occasions when marks should have been adjusted but, because the teacher was well-respected or advocated strongly, no changes were made. Some, when changes were suggested, requested no changes be made because of repercussions at school. Such concerns led to the demise of group moderation in the SACE.

Earlier (B. Atherton, 2009), I investigated teachers' understandings of, and attitudes towards, social moderation with teachers of Studies of Societies. Teachers were “generally satisfied” (p. 183) with the process with many commenting that an examination was inappropriate because the skills taught in the subject could not be validly assessed by examination. All participants saw social moderation as necessary “to ensure consistency across the state” (p. 110) and acknowledged its educative benefit.

## **1.9 Research goals**

The SACE assessment changes in 2011 were in response to the SACE Review (Crafter et al., 2006) which “proposed to strengthen the quality assurance processes ... to one that promotes greater teacher involvement” (p. 125).

The current research was based on the premise that teacher learning from the

processes would necessitate involvement in those processes. Consequently, the first goal of the current research was to explore the extent of teacher involvement in the SACE Board's quality management processes in 2011 and the possible association between involvement and teacher characteristics such as experience, sector or location.

The last fifteen years has seen much research into professional learning for teachers (e.g., Desimone, 2009; Garet et al., 1999; Hawley & Valli, 1999; Loucks-Horsley, Love, Stiles, Mundry, & Hewson, 2003; Timperley, Wilson, Barrar, & Fung, 2007; Wei, Darling-Hammond, Andree, Richardson, & Orphanos, 2009), which has identified factors contributing to the effectiveness of teacher learning. The second goal of the current research was to analyse the extent to which the forms of the SACE Board's quality management processes were consistent with those described in the literature as contributing to effective teacher learning.

Among researchers, a common finding regarding the effectiveness of teachers' professional development activities related to their content (e.g., Desimone, 2009; Garet, Porter, Desimone, Birman, & Yoon, 2001; Timperley, 2008; Wei et al., 2009). The third goal of the current research was to examine the content of the SACE Board's quality management processes for evidence that they might contribute to teacher learning.

While the first three goals examined the SACE Board's quality management processes per se, the fourth sought evidence of teacher change in the area of assessment that might have occurred through involvement in the SACE Board's quality management processes. The research also explored possible associations between any learning and teacher factors such as experience, location and involvement in the quality management processes.

## **1.10 Chapter summary and thesis outline**

This chapter has described the increased use of SBA in senior secondary assessment and the quality management that accompanies its use in this high-stakes environment. The terms used in this thesis for the components of quality management were defined and literature discussed that cites the educative value of social moderation. The SACE Board quality management processes were outlined,

the researcher's position described and the research goals defined.

These goals required the collection of data on teacher involvement, the form of the quality management processes, the content of the quality management processes and evidence of teacher learning in the area of assessment. With such a range of data sought, it was believed that multiple sources were required. The next chapter describes how the existing literature, particularly that associated with teacher learning and SBA, was scrutinised and used in developing the research approach. The evolution of the four goals into the ultimate research questions is described.

Chapter 3 explains the methodology of the current research: the methods used for data collection and the approaches taken in data analysis. It also presents the ethical issues associated with the research and explains why the final research varied from the original concept. In Chapter 4 the participants are described and teacher involvement in the quality management processes analysed. Chapters 5 and 6 examine the form and content of the quality management processes, while Chapter 7 examines possible teacher learning. The results are then discussed in Chapter 8. Chapter 9 summarises the findings, reflects on the research process and makes suggestions for further research.

## CHAPTER 2: THE RESEARCH CONTEXT AND LITERATURE REVIEW

### 2.1 Introduction

The main aim of this thesis was to explore whether, and, if so how, quality management processes associated with senior secondary school-based assessment (SBA) might contribute to teacher learning. Analysis of the quality management processes considered a number of strands in literature: effective professional development activities, other forms of teacher learning, adult learning and deep learning.

The reported research investigated the quality management processes of the South Australian Certificate of Education (SACE) which is considered to be *high-stakes*<sup>3</sup> because it might “significantly affect the immediate and sometimes long term life-course and life-chances of students” (Crafter et al., 2006, p. 126). The assessments are used “as the basis for awarding qualifications” (Hipkins & Robertson, 2011, p. 7) and for calculating Australian Tertiary Admissions Ranks (ATAR), used for tertiary admission (SATAAC, 2012). The research undertaken was in connection with the academic year 2011 in which all subjects implemented a 70% weighting of SBA, performance standards and social moderation of SBA. The latter two directly affected teachers’ assessment practices. It was considered that teachers’ responses to these changes might be more evident in the first year of the new system. Thus 2011 presented a unique opportunity to explore the potential for teacher learning in quality management processes.

Four areas of research were identified for this thesis: teacher involvement in the quality management processes, the form of the processes, the content of the processes and changes in teachers’ assessment beliefs and practices that might be

---

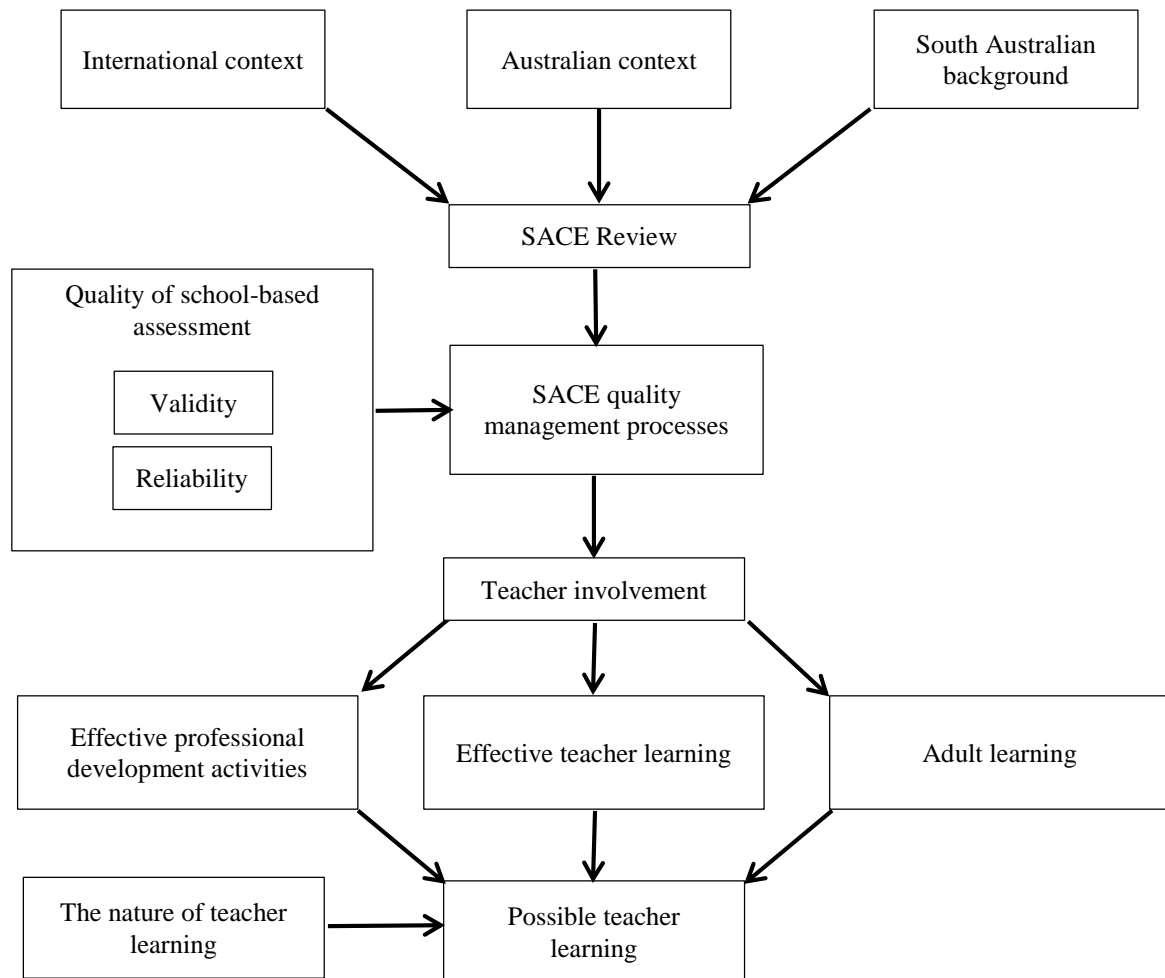
<sup>3</sup> In this thesis, the term *high-stakes* refers to end-of-secondary assessment although it has been argued that the term is also applicable to other assessments. For example, Connolly et al. (2012) argued that, with the introduction of national testing, assessment in the first eleven years of schooling earlier years had become high-stakes, while Smail (2013) believed the introduction of National Standards in New Zealand had “undoubtedly raised the assessment stakes for schools” (p. 11).



associated with the quality management processes.

This chapter begins by examining the type of teacher learning that might be associated with quality management processes: what is meant by learning generally and in the particular context of teacher learning.

The chapter then describes the research context, starting with an international view, then focussing on the Australian and South Australian scenes. This is considered necessary because current research explored whether teachers had been informed of the evidence used by the SACE Review Panel (Crafter et al., 2006) in making their recommendations. The literature search process is then described. Existing learning theories are examined for their applicability to analysing the quality management processes for potential teacher learning. Gaps in existing knowledge are identified and used to generate sub-questions, thus clarifying the domain of the current research.



**Figure 1: Relationship between concepts addressed in literature review**

The chapter concludes with a summary of the research questions and sub-questions.

## 2.2 Possible teacher learning

This section examines aspects of teacher learning of relevance to the current research. First, it examines the time required for teacher learning, followed by an analysis of the various ways in which teacher learning might be characterised, thus providing a basis for examining teacher learning that might be associated with the quality management processes.

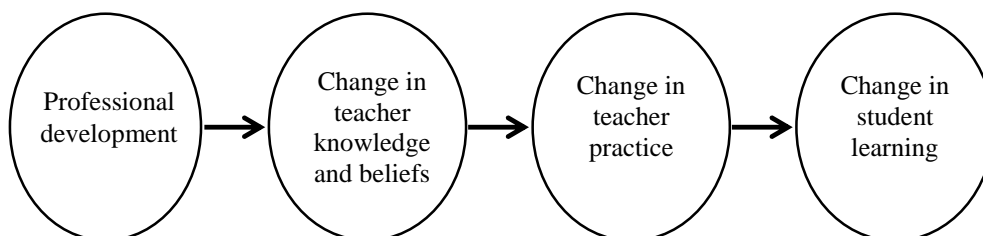
Concerning the time required for teacher learning, Guskey (2000) cautioned that any measurement of changes in teacher practice should be made after some time “to allow participants to adapt the new ideas and practices to their setting” (p. 85). Time is needed for teachers to fully implement changes in teaching practice (Hipkins & Robertson, 2011; Loucks-Horsley et al., 2003; Wylie & Hodgen, 2010). Conducting

the current research over a twelve-month period allowed sufficient time for some evolution of teachers' practices and beliefs during the research period.

The literature was examined for models for identifying and describing teacher learning that might be associated with the quality management processes. Some (e.g., Kuijpers, Houtveen, & Wubbels, 2010; Lovett & Gilmore, 2003) were "context-specific" (Evans, 2014, p. 182) and considered inappropriate for the context of the current research. There were, however, several models providing a framework for examining teacher learning associated with involvement in the quality management processes.

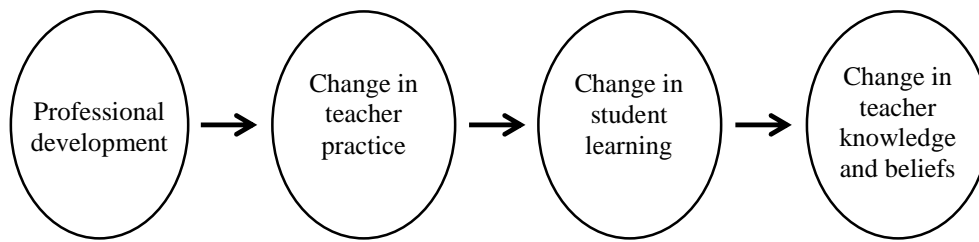
While behaviourists describe learning as a change in observable behaviour as a result of experience (James & Pedder, 2006b), others describe it as a change in "behavior potential" (Kagan, Segal, Havemann, Baucum, & Smith, 2004, p. 139). In the matter of teacher learning, some authors focus on changes in teacher practice (Postholm, 2012; Wei et al., 2009), while others mention changes in their beliefs (Desimone, 2009; Guskey, 2000). Whether changes in teacher practice precede or follow changes in beliefs is contentious (Richardson & Placier, 2001). James and Pedder (2006a) commented on the "long-running debate about whether it is necessary to change beliefs and values in order to change practices, and which should come first" (p. 111).

The "conventional wisdom" (Loucks-Horsley et al., 2003) was that professional development activities should address teachers' beliefs (as shown in Figure 2), leading to changes in teaching practices.



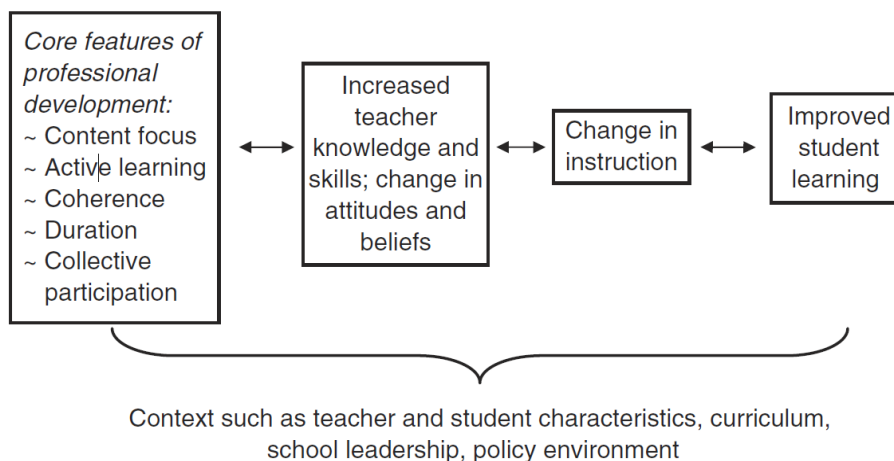
**Figure 2: Traditional model: professional development through changing teacher knowledge**

Guskey (1986) proposed that changes in teacher practice *preceded* changes in their beliefs, caused by seeing benefits for students, as shown in Figure 3.



**Figure 3: Guskey's model: professional development through changing teacher practice**

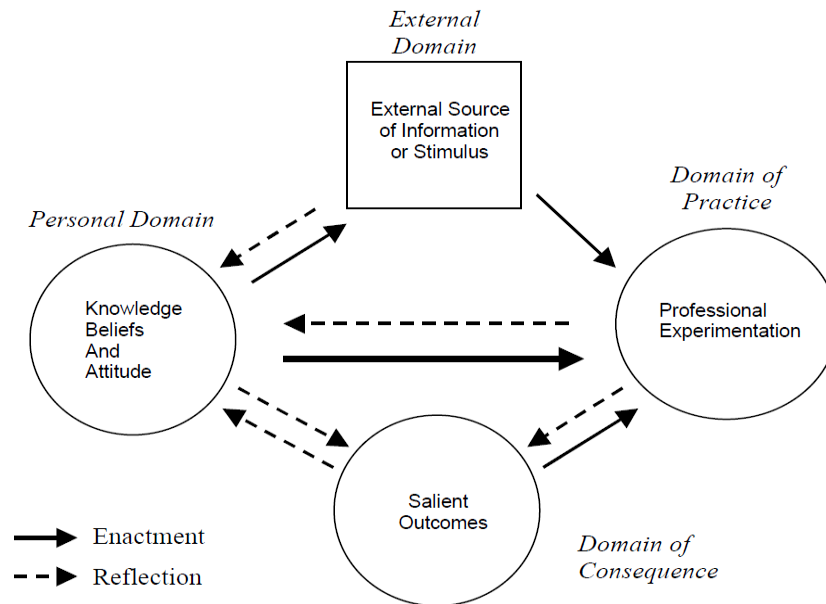
Huberman believed that the process of teacher change was more likely to be cyclical than linear (reported in Guskey, 2002) and changes in attitudes, beliefs and practices were described as occurring “in a mutually interactive process” (Loucks-Horsley et al., 2003, p. 48). Desimone (2009), for example, proposed a model (Figure 4) showing the same order as the traditional model (Figure 2) with two-way arrows indicating the “interactive, nonrecursive relationships” (p. 184) between the components.



**Figure 4: Desimone's model for the effects of professional development**

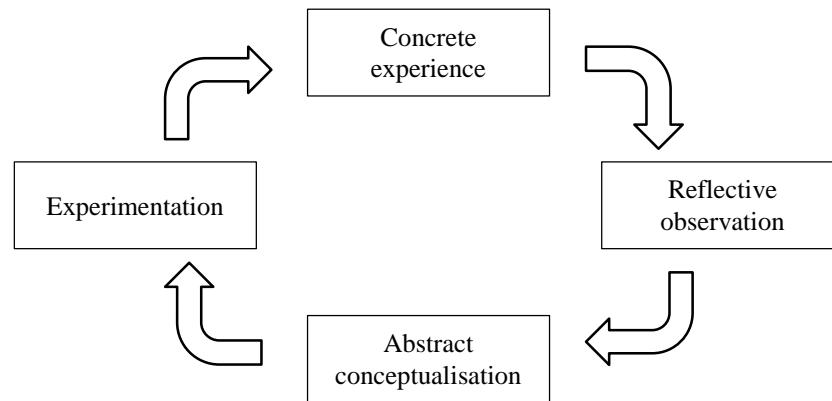
It is noted that the above models all incorporate change in student learning outcomes. Evans (2014) suggested that attention to student learning as the outcome of teacher learning is an American phenomenon and that elsewhere the professional development of teachers is considered “a justifiable end in itself” (p. 181). The current research focussed on teacher learning rather than student learning. In particular, the current research sought evidence of changes in teachers' attitudes, beliefs and practices in the area of assessment, which might be associated with involvement in the quality management processes.

Clarke and Hollingsworth (2002) described an *interconnected model* (Figure 5) that moved away from the linear arrangements of the above models, showing “the complexity of professional growth through the identification of multiple growth pathways” (p. 950).



**Figure 5: Clarke and Hollingsworth’s (2002) interconnected model of teacher learning**

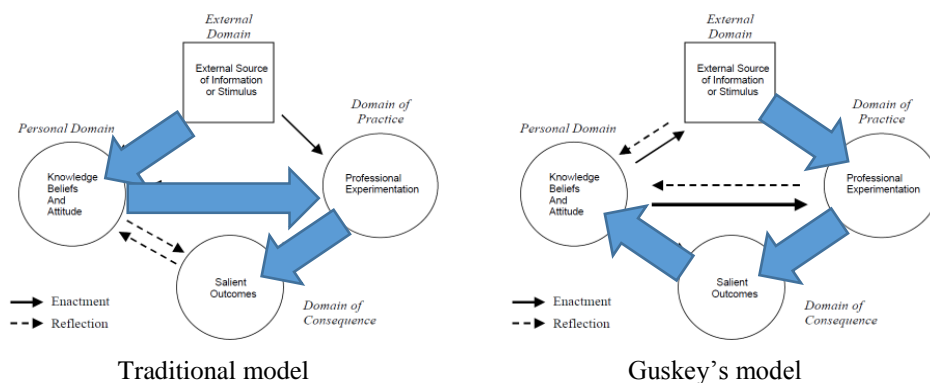
In this model, the Personal, Practice and Consequence domains constitute the personal world of the teacher, distinct from the external domain. Changes occur through teacher *enactment* “of something a teacher knows, believes or has experienced” (Teacher Professional Growth Consortium cited in Hollingsworth, 1999, p. 129) and *reflection*, with the *salient outcomes* being new teacher conclusions, based on reflection on new knowledge or new practice. The importance of reflection in the learning process appears to comply with the understandings of *experiential learning* (Kolb, 1984) and *reflective practice* (Schon, 1987). Figure 6 shows a simplified version of the experiential learning cycle of Kolb (1984).



**Figure 6: Experiential learning cycle based on Kolb (1984)**

In Clarke and Hollingsworth’s (2002) interconnected model of teacher learning, the salient outcomes, those valued by teachers, depend “entirely upon the value system of the teacher” (p. 954). In an earlier version (D. J. Clarke & Peter, 1993), the “domain of inference” consisted of “valued” outcomes (p. 170).

Clarke and Hollingsworth (2002) showed that it incorporates the traditional (Figure 2) and Guskey models (Figure 3) as shown in Figure 7 (based on D. J. Clarke & Hollingsworth, 2002, pp. 960, 961).



**Figure 7: Traditional and Guskey learning models incorporated in the interconnected model (Clarke and Hollingsworth, 2002)**

Clarke and Hollingsworth (2002) specified that change in practice is not restricted to the classroom and that external stimuli are not restricted to traditional forms of professional development. These points were relevant to the current research because changes in the area of assessment and diverse “external stimuli” were available to teachers. The authors indicated that the model recognises that teachers learn in different ways. The specified “key shift ... from programs that change teachers to teachers as active learners” (p. 948) is addressed in two ways in the

current research: whether active learning was encouraged in the quality management processes (see 2.6.2.2) and the extent to which teachers adopted active or passive roles as learners.

The model allows for teacher growth as “evolving practices” (p. 955) in an iterative process of enactment and reflection. A change in one domain may, or may not, lead to change in another. Changes connected by reflection or enactment are called *change sequences*, which may be “fleeting ... [and] quickly relinquished” (p. 958). Change of a more permanent nature is considered as *growth* rather than *change*. It was considered that less permanent instances of change might be identifiable in the current research, which was restricted to twelve months. Although not initiating their use, Clarke and Hollingsworth used the terms *constraints* and *affordances*, which will be used in this thesis for factors that appeared to hinder or facilitate teachers learning from involvement in the quality management processes.

The concept of teacher change occurring incrementally was also used by Evans (2014) who referred to “micro-level” professional development as “individual, singular ‘episodes’” (p. 186) that are parts of a “bigger picture” of teacher learning. Evans (2014) suggested that one change episode may lead to another in “chain-reaction episodes” (p. 194) and considered her “micro-level development” analogous to Clarke and Hollingsworth’s (2002) change sequences. Evans (2014) argued that such episodes comprise a teacher accepting, consciously or unconsciously, a “better way of doing things” (p. 185). She argued, however, that “Change in knowledge, beliefs or attitudes does not necessarily involve reflection; more often it occurs spontaneously, or it evolves imperceptibly over time” (p. 185). She argued that teacher implementation of imposed changes not considered “better” did not constitute professional learning, “which requires intellectual or attitudinal development” (p. 192). The identification of change at the micro-level offered opportunities for identifying teacher change in the current research in one year, rather than the three to five years required for successful implementation of teacher change (Loucks-Horsley et al., 2003).

Thus, in the current research, possible teacher learning in the area of assessment was examined from two perspectives: whether changes in belief followed mandated changes in practice and whether changes might be described as change sequences

involving enactment and reflection (D. J. Clarke & Hollingsworth, 2002).

## **2.3 The context of the research**

The SACE assessment changes of 2011 did not arise in isolation. Although, for participants, the local setting was all-important, an examination of the broader context is desirable since it was considered by the authors of the SACE Review (Crafter et al., 2006). Some participants had experience of other systems and all potential participants would have been aware of changes on the national education stage. The context is described below, beginning with the international scene, then narrowing the focus to Australia and South Australia.

### **2.3.1 The international context**

The new SACE was a product of previous South Australian practices and the SACE Review Panel's (Crafter et al., 2006) interpretation of international and national trends. Two phenomena had been observed in many educational systems: increased reliance on teacher judgement in end-of-secondary assessment and the increased use of performance standards. While these are related, they will be examined separately. Subsequent sections describe how assessment changed for SACE teachers.

#### **2.3.1.1 *Increased use of school-based assessment***

A number of authors acknowledge the increased use of SBA in high-stakes assessment (Briseid & Caillods, 2004; Crisp, 2013; Davison, 2004; S. Johnson, 2013; Matters, 2006; Mercurio, 2008; Wolf, 1995), but suggest different rationales. One position is that it is reasonable for teachers to use their knowledge of students in the assessment of those students (Davison, 2004; S. Johnson, 2013): "Why rely on an out-of-focus snapshot taken by a total stranger?" (Davison, 2004, p. 306) Teachers are "closest to the action of student learning" (Crafter et al., 2006) and can best integrate assessment into student learning. Three other reasons, however, are more commonly cited: the changing role of senior secondary education (Briseid & Caillods, 2004; Mercurio, 2008), the impact of assessment on learning (Broadfoot, 2005; Harlen, 2006; 1995) and the relative qualities of school-based and external assessment (Black, 1998; Wolf, 1995). Each is now discussed.

The use of SBA was seen to be more appropriate as senior secondary education



served “an increasingly diverse group of students” (Mercurio, 2008, p. 2). As the age of compulsory education was raised, “pressures arose to provide school leaving certification ... for all” (Black, 1998, p. 14). Secondary education prepared students, not only for further learning, but also for work and participation in a democratic society (Briseid & Caillods, 2004). Since not all senior secondary students aspire to tertiary studies, external examinations, reflecting the assessment of universities (Crafter et al., 2006; Young, 1998), were considered inappropriate. Today many countries use a blend of SBA with external assessment (Briseid & Caillods, 2004; Mercurio, 2008, p. 11).

Increased use of SBA was also in response to research showing the possible negative impact of assessment on teaching and learning (e.g., ARG, 2006; Black & Wiliam, 1998; Harlen, 2004b; S. Johnson, 2013; Wolf, 1995). While Black and Wiliam (1998) reported that appropriate assessment could enhance student learning, others reported damage that assessment might have on learning. Harlen and Deakin Crick (2002) reported negative effects of assessment on teachers’ teaching and assessment practices and on the self-esteem of lower-achieving students.

The third factor contributing to increasing use of SBA was recognition of examinations’ limitations. An examination has a limited capacity to assess the teaching program (Black, 1998). Questions were raised about examinations’ *validity* and *reliability*, the terms commonly used to describe the quality of assessment. Proponents of SBA argue its better validity.<sup>4</sup> Because teacher judgements can be made more often (Maxwell, 2006; Wiliam, 2001) and with more task variation than examinations (Crisp, 2013; Long, 2006; Matters, 2006), SBA can assess a wider range of learning outcomes (ARG, 2006; Crisp, 2013; S. Johnson, 2013; Maxwell, 2006). Varied assessment tasks, possible with SBA, are needed because the multidimensional nature of learning “cannot be adequately measured by one instrument” (Suurtamm, Koch, & Arden, 2010, p. 400). The use of more assessment tasks allows for more comprehensive coverage of the learning domain, providing richer evidence of the qualities being assessed; the assessment is hence more valid.

---

<sup>4</sup> *Validity* refers to “how well the result really reflects the skill, knowledge, attitude or other quality it was intended to assess” (Harlen, 1994, p. 2).

It is also argued that SBA, with more tasks, may improve reliability<sup>5</sup> (Wiliam, 2001). With more opportunities for assessment, the effects of random variations that might occur in one-off assessments of student learning are reduced. The assessment is more reliable with more assessment opportunities; “consistently repeatable results” (Harlen, 1994) are not possible with one-off tasks. Thus SBA might improve both the validity and reliability of assessment.

There are, however, caveats associated with the use of SBA. It is costly to implement because of professional development, exemplar production and the need for moderation (Zepke et al., 2005). It is viewed as having “high validity but questionable reliability” (Wyatt-Smith, Klenowski, & Gunn, 2010, p. 61). Grading differences between schools and teachers have been acknowledged (Willingham, Pollack, & Lewis, 2002). Harlen (2005) reported bias and lack of reliability in SBA, with teachers influenced by student characteristics such as gender, social class and appearance (Black, 1998; Wyatt-Smith, 1999) and by prior impressions of the students and characteristics such as handwriting (Crisp, 2010a). Concerns over the reliability of SBA become more pronounced with high-stakes assessment (Crooks, Kane, & Cohen, 1996; Maxwell, 2006). Consequently, whenever SBA is incorporated into high-stakes assessment, quality management processes have been put in place (Mercurio, 2008).

Thus three reasons have been commonly presented for increased use of SBA: the changing role of senior secondary education, the impact of assessment on learning and the greater validity of SBA. The current research investigated the SACE Board’s quality management processes for evidence that they might contribute to the development of teachers’ assessment literacies (Crafter et al., 2006; Kilvert & Mercurio, 2007; Mercurio, 2008) by informing them of the reasons for the changes.

### **2.3.1.2 Increased use of standards-based assessment**

In the new SACE, teachers were required to assign grades to students through the use

---

<sup>5</sup> *Reliability* of assessment refers to the extent to which it is able to generate consistently repeatable results (Harlen, 1994).

of performance standards. Previously, Physics teachers had provided SSABSA<sup>6</sup> with a list of students in rank order, each with a mark out of 100. The distribution of marks was important, not the actual values, since the marks were statistically moderated. Grades were assigned by SSABSA. In the new system teachers were required to submit grades, absolute values rather than relative, for all students.

Because performance standards were new for most teachers, they were an important area for teacher learning in the new SACE. It is desirable, therefore, to examine the increasing use of standards-based assessment (Adie, 2008; Bennett, Tognolini, & Pickering, 2012; Connolly et al., 2012; Davison, 2004; Smaill, 2013).

The use of performance standards has been described as “a dominant assessment paradigm in English speaking countries” (Rawlins et al., 2005, p. 108; Zepke et al., 2005, pp. 6, 14). The OECD (2005), however, mentioned that several European countries were working on performance standards and Klenowski and Wyatt-Smith (2010b) mentioned their introduction into Germany and Norway. It seems that their use is not restricted to the Anglophone world. Klenowski and Wyatt-Smith (2010) described the development of national standards by many countries as a response to widely-publicised international assessments such as the Programme for International Student Assessment (PISA) or Trends in International Mathematics and Science Study (TIMSS).

Many reasons are given for the proliferation of standards-based assessment. Smaill (2012) and Adie (2008) described its dual purpose: addressing the learning needs of students, as well as serving system accountability requirements. Standards provide guidance to teachers in judging student work (Wyatt-Smith & Klenowski, 2013), requiring them to “investigate the meaning of the work that students generate” (Hipkins & Hodgen, 2011, p. 2). They are also seen to support “the progressive development of deeper knowledge and understanding, and the development of higher order thinking skills” (Adie, 2008, p. 2). By providing guidance to teachers, performance standards might help counteract the negative views of SBA mentioned above.

---

<sup>6</sup> The Senior Secondary Board of South Australia (SSABSA) was responsible for senior secondary assessment in South Australia prior to the SACE Board (Mercurio, 2003).

Performance standards require teachers to view student work in a different way. Wiliam (1992) argued that “Any assessment scheme works by comparing the performance of an individual with something else” (p. 17). The “something else” might be other students—*norm-referenced* assessment—or some external descriptor—*criterion-referenced* assessment. The shift from norm-referenced to criterion-referenced assessment is now examined.

### **2.3.1.2.1 Norm-referenced and criterion-referenced assessment**

Norm-referencing involves “essentially ... comparing a pupil's attainment with that of others” (Brown, 1980, p. 1) undertaking the same assessment. This means that a student's outcome “depends as much on the performance of others as it does on his or her own” (Wiliam, 1992, p. 17), with no information provided about what a student has achieved (Brown, 1980). Norm-referenced assessment is common when students undertake a common task such as an examination.

The comparison of the SBA of different teachers who use different tasks requires some method of *linking* (Linn, 1993) the student outcomes. This may be achieved statistically, known as *statistical moderation*, using marks from a common task, such as an examination, for comparing groups of students. Alternatively, the achievement of students from different cohorts may be compared using a set of common criteria, *criterion-referenced* (Brown, 1980) or *standards-referenced assessment* (Forster, 2007). The linking between different teachers' judgements is thus undertaken using professional judgement.

In criterion-referenced assessment student performance is described in terms of “what the pupils know or can do, without reference to the performance of others” (Brown, 1980, p. vii). Instead, the performance is compared to “some well-defined object” (Wiliam, 1992, p. 17). Criteria may describe performances at a grade boundary (Crisp, 2010b) or apply to “typical” students within a grade (e.g., Board of Studies NSW, 2009).

Each system defines its own terms. Klenowski and Wyatt-Smith (2010b) wrote that “the word *standard* is ubiquitous yet difficult to define” (p. 109), but made a distinction between content and achievement standards; content standards are “the knowledge and/or processes that are taught” (p. 110) while achievement standards

apply to student performances. Content standards are commonly prescribed in a syllabus describing what is to be taught. In systems with standards-based assessment, syllabuses often include achievement standards, also known as *performance standards*. Performance standards were new to the SACE in 2011 and a focus of teacher learning in the current research. Thus, in this thesis the term *standards* will refer to *performance standards* rather than *content standards*.

### **2.3.1.2.2 Performance standards**

Performance standards serve as a basis for “the progressive development of deeper knowledge and understanding ... [and] promote the transparency of the curriculum and assessment practices” (Adie, 2008, p. 2). They provide teachers with a language they can use in “articulating their understandings and beliefs” (Adie, 2008, p. 10). Teacher acceptance of performance standards depends, however, on their being understandable and credible (Hargreaves, Earl, Moore, & Manning, 2001). They need to be “so clear and unambiguous that reliable, parallel assessments can be derived” (Wolf, 1995, p. 3). Involving subject teachers in their development would assist in their reliable use (Harlen, 2007) and teacher appreciation of their relevance (Brown, 1980).

Performance standards might be written from the curriculum or derived from analysis of existing student work (Forster, 2007), with the latter preferred (Burt & Stapleton, 2010; Forster, 2007; QSA, 2010e). Teacher knowledge of previous student performance in writing performance standards creates a paradox: the performance standards arise from evidence gathered through norm-referenced assessment (Brown, 1980).

Performance standards vary in the degree of detail they provide. Those that are too general give “too little idea of what evidence is needed and how it would be judged” (Harlen, 2007, p. 143). Although some degree of detail is desirable, overly detailed criteria become “too numerous to be manageable” (Harlen, 2007, p. 142). Sadler (2007) described this extreme as consisting of criteria “reduced to pea-sized bits to be swallowed one at a time” (p. 390). He described how teachers may not be able to specify in advance which criteria are to be applied to any particular response and suggested that teachers needed freedom to select from a “large pool of potential criteria that could legitimately be brought to bear” (Sadler, 1989, p. 124).

If performance standards are to be understandable to teachers, teachers should be involved in their creation from a basis of student work. They need to be neither too prescriptive nor too general. They should describe a development in understanding or skills. Writing of the UK context, where they serve as “developmental standards” (Maxwell, 2009), Harlen (2007) stated that performance standards should not be task specific if they are to provide the best help for reliable teacher assessment. Written standards can, however, serve other purposes and may be applicable to individual assessment tasks or to a collection of tasks, requiring different forms of rubric (Maxwell, 2009). Thus performance standards in Australian states differ in their degree of detail and whether they apply to individual tasks, assessment types or overall judgement of student work.

No matter how well-written, performance standards are always open to interpretation (Crafter et al., 2006) and require qualitative teacher judgement (Klenowski & Wyatt-Smith, 2010b). Teachers use other factors, however, when assessing student work, even when provided with performance standards (Adie, 2008; Cooksey, Freebody, & Wyatt-Smith, 2007); “criteria statements represent a necessary, though not sufficient, condition for achieving comparable grading outcomes” (Wyatt-Smith, 1999, p. 219). Teachers’ interpretations of performance standards descriptors are affected by teachers’ backgrounds and experiences (Burt & Stapleton, 2010). Teachers need to internalise the criteria and the standards if they are to arrive at judgements “comparable with other teachers” (Daugherty, 1997, p. 33). Several approaches are advocated for addressing this need in the use of performance standards for SBA: exemplars of student work, professional development and discussion among teachers.

The use of exemplars has been advocated for SBA (Harlen, 1994; Heldsinger & Humphry, 2013; Maxwell, 2006; Sadler, 1987) to provide teachers with “reference points against which comparisons can be made [thereby helping] ... to fix the standard” (Crisp, 2010a, p. 53). Sadler (1987) indicated that many exemplars are needed if there are many criteria. Baird (2007), however, found that examiners, deliberately given a mislabelled exemplar or no exemplar, were still able to set standards comparable with other markers through reference to “their own mental models of the standard” (p. 142). Since she was working with experienced

examination markers, it would be imprudent to make similar inferences for subject teachers using SBA.

“Teachers need both initial training and continuing opportunities to share judgments for them to become familiar with, and at ease with, operationalising the procedures, the criteria and the standards applicable” (Daugherty, 1997, p. 33). The ARG identified pre-service and in-service professional development as “key requirements” (ARG, 2006, p. 1) for the implementation of SBA. The SACE Board offered professional development for teachers of the new SACE (SACE Board of South Australia, 2010i).

The opportunity for teacher discussion builds “a critical and analytical understanding of the standards and the ability to interpret and apply them” (Crafter et al., 2006, p. 139) and provides teachers with opportunities to “reflect on the rationale for their judgement decisions” (Adie, 2008, p. 4), “articulating their understandings and beliefs” (Adie, 2008, p. 10) as they explain their decisions. Daugherty (1997) described the need for continuing discussion opportunities in order to become familiar with the criteria and the standards. Performance standards *per se* do not provide the common understanding among teachers.

When a changed assessment system was introduced in New Zealand, Black (reported in Starkey et al., 2009) argued that professional development should focus on broad issues such as the principles underpinning standards-based assessment. This was endorsed by Starkey et al. (2009) who stated that “effective professional development must engage teachers with the “bigger” ideas ... not simply teach operational skills” (p. 182).

The current research investigated the extent to which the SACE Board quality management processes facilitated teacher understanding and use of performance standards. Had information been provided about the “bigger” ideas such as their purposes, advantages and disadvantages?

### **2.3.2 The Australian context**

Although participants may not have been aware of the international context, they were aware of the changes in the SACE and, in interviews, approximately half of the

participants mentioned other states' assessment systems and the national curriculum. Because awareness of pending national changes may have affected willingness to commit to the SACE changes, it is pertinent to describe the Australian scene.

Although Australia became a nation in 1901, it maintains eight educational systems (Matters 2006; Mercurio, 2008), six states and two territories,<sup>7</sup> with parallel traditions of public examinations (Mercurio, 2008).

In the federal system, education is traditionally the responsibility of the states (A. Reid, 2009), although this is not specified in the constitution (Cumming, 2010, p. 421). Differences between the states emerged gradually, and recently the state-federal balance has shifted as the Federal government has used “the power of the purse” (Cumming, 2010, p. 419) to impose conditions on the states (Ferrari, 2007; Klenowski & Wyatt-Smith, 2010b; A. Reid, 2009).

### **2.3.2.1 School-based assessment in Australia**

Australia has an international reputation for its use of SBA in high-stakes environments (Cumming, 2010) with all states incorporating teacher judgements into end-of-secondary assessment (Cumming & Maxwell, 2004; Matters, 2006; Mercurio, 2008). The weighting assigned to the SBA, however, varies, with only Queensland and the ACT using it solely.

Differences between the states began to appear when the increasing number of students, with differing needs, necessitated a changed role for senior secondary education. While other states continued with an examination-based system, Queensland abandoned this aspect of its British heritage (Mercurio, 2008) some forty years ago (QSA, 2010b). In Queensland “public outcry over the exposure of poor examination practices ... catalysed existing public concern” (Maxwell & Cumming, 2011, p. 189) which sought a secondary system more responsive to students' changing needs. A review of the examination system generated the Radford Report, which recommended sweeping changes in curriculum and assessment (Maxwell & Cumming, 2011). Queensland's adoption of a system using only SBA occurred after

---

<sup>7</sup> Because the territories are educationally independent the term *state* will be taken to include the two territories as well as the six states.



the 1972 examinations (Maxwell & Cumming, 2011), with similar changes following in 1976 in the ACT (Mercurio, 2008). Although Maxwell and Cumming (2011) expressed surprise that only the ACT followed Queensland's path, this could be due to the conviction of each state that its system—be it education, railway or transport—is superior.

The extent to which teachers value moderation is influenced by their beliefs and attitudes (Connolly et al., 2012; Klenowski & Adie, 2009). My participants' beliefs and attitudes might have been influenced by experience of the old SACE and awareness of other systems, particularly regarding the forms of moderation and ATAR calculations. Four participants had taught in other states, but others referred to other states, particularly Queensland, cited in the SACE review (Crafter et al., 2006) as worthy of emulation.

While all states use SBA, they use different means of quality control; that is, different methods for adjusting students' marks. Apart from South Australia, the use of social moderation is limited to Queensland (QSA, 2010d) and the Australian Capital Territory (ACT Board of Senior Secondary Studies, 2012). In Queensland and the ACT, tertiary entrance scores are determined after statistical adjustment of student results based on the Queensland Core Skills Test (QSA, 2010b) or the ACT Scaling Test (ACT Board of Senior Secondary Studies, 2012). In other states SBA results are statistically moderated against examinations (Matters, 2006), although Victoria also uses its General Achievement Test (VCAA, 2013). South Australia undertakes no inter-school statistical adjustments of moderated grades prior to tertiary entrance calculations, so SACE moderation outcomes might be considered of higher stakes than in other states.

### **2.3.2.2 Performance standards in Australia**

Performance standards, used in all states, are usually written as a grid, sometimes referred to as a rubric (e.g., DEECD, 2007). The various states, however, use different terms. While the SACE calls them performance standards (SACE Board of South Australia, 2011d), they are also known as *performance descriptors* (VCAA, 2008), *performance band descriptions* (Board of Studies NSW, 2009) and *standards* (QSA, 2007a).

Performance standards in the various states also differ in their degree of detail, the names used for their components and whether they apply to individual tasks, assessment types or overall judgement of student work. It is not considered necessary to detail the differences between the states. Because the current research involved the new South Australian curriculum, the new SACE terms to be used in this paper are described later.

### **2.3.3 The South Australian scene**

Motivation affects the amount of time people are willing to commit to learning (Bransford, Brown, & Cocking, 1999). For the current research, it was considered that previous experience with SACE assessment might influence teachers' attitudes towards assessment reforms. This section describes the changes made in the SACE to provide the context for teacher attitudes reported later. It begins by describing the justification given for the changes, then reports on the processes used to write the performance standards. This is followed by an overview of the quality management processes and the ways in which teachers might be involved in those processes.

#### **2.3.3.1 *School-based assessment***

In South Australia, SBA had been incorporated into senior secondary for many years. Around the time of Queensland's Radford report, the South Australian Education Department introduced a Secondary School Certificate as an alternative for students not wishing to continue to university (Mercurio, 2003). Previously, through its control of tertiary entrance examinations, the University of Adelaide had effectively dictated the South Australian secondary curriculum (Mercurio, 2008). The Department perceived the need for an alternative certificate for students who were completing 'what was increasingly becoming a societal expectation of all students—five years of secondary education' (Mercurio, 2003, p. 128). The Department wanted schools to reflect their students' individual needs and not have all subjects publicly-examined. The outcome, however, was that the matriculation curriculum was seen as superior to other subjects. Thus began a battle for "parity of esteem" (Mercurio, 2003, p. 128) between subjects which served different needs, played out between 1969 and 1983, but influencing the present century as universities allowed only some subjects to be used for university selection (Crafter et al., 2006; Mercurio, 2003). One significant objective of the SACE Review (Crafter et al., 2006) was to

dispense with the different assessment methods in the old SACE, which contributed to “hierarchies of knowledge” (p. 130); a common system was recommended for all subjects. The current research explored whether teachers had been informed of the reasons for the changes, thus contributing to assessment literacy (Kilvert & Mercurio, 2007).

The SACE Review also reported a “decreasing reliance on teacher judgement” (Crafter et al., 2006, p. 129) with a trend to focus on adjustment of teachers’ results rather than expending effort on assisting teachers in making their judgements; quality control was preferred to quality assurance. The Review (Crafter et al., 2006) recommended “greater involvement of teachers in moderation processes” (p. 130).

For Physics teachers, three significant changes were implemented in 2011. The weighting of SBA was increased from 50% (SACE Board of South Australia, 2009) to 70% (SACE Board of South Australia, 2011d), to be common to all subjects.<sup>8</sup> Social moderation (SACE Board of South Australia, 2011g) replaced statistical moderation (SACE Board of South Australia, 2010j) and performance standards were introduced for student assessment (SACE Board of South Australia, 2011d).

For Physics teachers, the increased reliance on teacher judgement was to be evident in the increased weighting of SBA and greater teacher involvement in moderation. The SACE Review Panel saw both as desirable (Crafter et al., 2006) and SACE Board executive officers endorsed greater teacher involvement (Kilvert & Mercurio, 2007). The *future* SACE Office, responsible for implementing the changes, “*proposed* [italics in original] to encourage and support participation by teachers on the final moderation and external assessment panels” (*future* SACE Office, 2008, p. 5).

An underlying premise of the current research was that, if teachers were to benefit from the quality management processes, they needed to be involved. Teachers engage in professional development to become better teachers (Guskey, 2002), commonly defined in terms of student success (Guskey, 2002; Pedder, James, &

---

<sup>8</sup> Except for the Extended Learning Initiative (later renamed Research Project) for which the recommendation was for 70 per cent (Crafter et al., 2006, p. 132).

MacBeath, 2005). A major professional development issue involves equal access for all teachers to quality professional development (Loucks-Horsley et al., 2003). The current research explored the extent of teacher involvement in the SACE Board quality management processes of the 2011 assessment cycle and whether this represented an increase over previous years. Reasons for teacher involvement or non-involvement were also explored so that means of improving teacher involvement might be proposed.

### **2.3.3.2 Performance standards in the SACE**

The adoption of performance standards represented an important change for SACE teachers. This section describes the performance standards, descriptions of their writing and their intended functions, providing a background to the experiences of participants in the current research.

The 2011 SACE Physics syllabus<sup>9</sup> (SACE Board of South Australia, 2011d) defined the performance standards as describing student performance at five grades (A–E) in thirteen *specific features*<sup>10</sup> (e.g., *Design of physics investigations* and *Demonstration of skills in individual and collaborative work*) grouped within four *assessment design criteria* (Investigation, Analysis and Evaluation, Application, and Knowledge and Understanding).

While some SACE subjects had used performance standards previously, Physics—the subject at the centre of the current research—had not. The 2011 Physics syllabus (SACE Board of South Australia, 2011d) included *learning requirements*, *assessment design criteria* and *specific features*, which corresponded to the *learning outcomes*, *criteria for judging performance* and *clarifying questions* of the previous year (SACE Board of South Australia, 2010l). Only the performance standards were entirely new, with no equivalent in previous documents.

Forster, having been contracted by the *future* SACE Office for advice on performance standards, reviewed their implementation elsewhere and recommended

---

<sup>9</sup> The term *syllabus* is used in this thesis although the term used by the SACE was *curriculum statement* in the old SACE and is *subject outline* in the new SACE.

<sup>10</sup> In 2013 the number of specific features was reduced to twelve.

“an iterative process beginning with examination tasks and elaborating the descriptions through reference to teacher developed assessments” (Forster, 2007, p. 7). She noted Western Australia’s experience of developing performance standards solely from curriculum documents had created “serious problems” (p. 7) and proposed that the performance standards should be tested during development in 2008.

The *future* SACE Office chose, however, to develop performance standards from curriculum documents rather than existing student work. Teachers in subject reference groups commented on the suggested performance standards, which were edited, circulated to schools and revised before implementation. At no stage were they viewed in the light of actual student work.<sup>11</sup> Consequently, teacher acceptance of the changes to the SACE may have been jeopardised by the manner of the performance standards’ development.

The SACE performance standards are “a central component of the curriculum and assessment framework” (Forster, 2007, p. 2) and “at the centre of the assessment and reporting processes” (SACE Board of South Australia, 2010k). Each syllabus (e.g., SACE Board of South Australia, 2011d, p. 83) specifies three uses of performance standards:

- during the teaching and learning program the teacher can give students feedback with reference to the performance standards
- students can use the performance standards to identify current performance and areas that need to be addressed for improvement
- at the completion of study the teacher refers to the standards in making a decision about the quality of the student’s learning.

Performance standards apply to assessment task design and for assigning grades at the assessment type level and for overall subject performance (SACE Board of South Australia, 2010k).

---

<sup>11</sup> Information on this process has been derived from my experience as a member of one of the subject reference groups, contact with three of the writers during the writing period and interview comments of five participants in the current research involved as writers or as members of subject reference groups.

Even with mandated change, teachers “have a considerable amount of discretion as to whether they implement the change in their classrooms” (Richardson & Placier, 2001, p. 909). Teacher commitment depends on their ability “to see the reasons for change, grasp the point of it, and be convinced it is feasible and will benefit their students” (Hargreaves et al., 2001, p. 118). The current research sought evidence that the quality management processes raised teacher awareness of the advantages and disadvantages of performance standards and equipped teachers in their use. It also gathered evidence of their use in task design and in grading student work.

### **2.3.3.3 SACE Quality management processes**

The SACE Board stated that quality management of its assessment processes was to “assure the community that assessment of learning in the SACE is fair, valid, and reliable” (SACE Board of South Australia, 2010f). The Board committed itself to quality management processes for both school-based and external assessment using a “four-phase process of planning, clarifying, confirming, and improving” (SACE Board of South Australia, 2010i, p. 1).

The four phases for SBA were described in *SACE Moderation Procedures* (SACE Board of South Australia, 2010i). *Planning* included the opportunity for teachers to attend planning forums and for them to submit, for approval, the school’s Learning and Assessment Plans (LAPs). In *clarifying* the Board provided forums where teachers could gain a “shared understanding of the performance standards and ... how to apply them consistently” (SACE Board of South Australia, 2010f). Thus *planning* and *clarifying* constitute the quality assurance component of quality management. In the third phase, *confirming*, the Board was to confirm that “the subject’s performance standards ... have been applied correctly and consistently across schools” (SACE Board of South Australia, 2010i, p. 1)—quality control of SBA. Confirming involves social moderation, where moderators, usually teachers, review samples of student work and adjust grades if the assessments “are inconsistent with the performance standards” (SACE Board of South Australia, 2010k). The fourth phase, *improving*, constitutes the quality review component of quality management; in this the SACE Board monitors and analyses the quality management processes. Subsequently it provides schools with data “to enable them to monitor and review each phase of their own quality assurance processes” (SACE Board of South

Australia, 2010f) and itself “uses the data and feedback from schools to improve its procedures” (SACE Board of South Australia, 2010f).

#### **2.3.3.4 Teacher involvement in quality management processes**

This section introduces the quality management processes of the 2011 SACE assessment cycle and the ways in which teachers might be involved.

*SACE Moderation Procedures* (SACE Board of South Australia, 2010i) identified four ways in which teachers might be considered as “clients” of quality management processes:

- attending a planning forum
- receiving feedback regarding the school’s LAP
- attending a clarifying forum
- receiving feedback from moderation.

Two of these, attending forums, conform to the common format of professional development activities with teachers as attendees. The other two—receiving feedback—involved the teachers as recipients of advice about their own assessment processes.

*Stage 2 Quality Assurance & Assessment Panels in 2011* (SACE Board of South Australia, 2010m) identified three additional processes in which teachers could be involved as assessment panel members:

- LAP approval
- clarifying forum support
- moderation.

Therefore, there were seven ways in which teachers might be involved in quality management processes, investigated in the current research for their potential for teacher learning. Thus the literature research sought the features of effective professional development activities and effective teacher learning, and explored what evidence might demonstrate teacher learning that might be associated with involvement in SACE Board quality management processes.

## 2.4 Empirical evidence for teacher learning from quality management processes

The first aspect of the literature search sought empirical evidence of learning associated with involvement in quality management processes. Professional learning from involvement in social moderation had been cited in the SACE Review (Crafter et al., 2006) as the reason for recommending the adoption of social moderation. The Review, however, cited no evidence in the body of the report to support the claims of the educative benefit of social moderation. The endnotes referred to two papers (Gipps, 1994b; Harlen, 1994), neither of which cited evidence either.

Clearly a more extensive search was necessary. The current research concerned the intersection of two fields—teacher professional learning and the quality management of SBA, particularly where that incorporated social moderation. From my role at SSABSA, where I had led professional development sessions in the use of SBA, and my earlier research (B. Atherton, 2009), I had knowledge of the literature associated with SBA and social moderation. The claim for teacher learning associated with social moderation was found in many sources (e.g., ARG, 2006; Harlen, 2004a; Matters, 2006). SACE Review panellists had visited Queensland (and other states) “to explore recent developments in curriculum and assessment” (Crafter et al., 2006, p. 29). The belief was expressed that most Queensland panellists “would see participation as a valuable professional development opportunity” (Maxwell, 2006, p. 6).

Although the current research considered the possibility of teacher learning from involvement in quality management processes *in toto*, the initial literature search focussed on social moderation since this formed the basis of the recommendations of the SACE Review (Crafter et al., 2006). Articles mentioned in the previous paragraph were scrutinised for supporting evidence. None referred to research evidence supporting the claims, suggesting a possible gap in the literature. Consequently, further searches were undertaken, first in peer-reviewed articles and then in other sources.



### **2.4.1 Peer-reviewed claims linking social moderation and teacher learning**

The first search was for peer-reviewed articles describing research into possible teacher learning associated with social moderation. Searches of the Education Resources Information Centre (ERIC) database using “Moderation + teacher learning” and “Moderation + professional development” yielded three articles.

The earliest (Malone, Long, & De Lucchi, 2004) described a Californian collaboration between four primary teachers who, using the same assessment item, collaboratively evaluated student work in a series of four two-hour meetings. Reflecting on the process, teachers reported improvement in subject knowledge as well as instruction and assessment skills. Although different from SACE Board processes, the research supported the possible professional benefit of collaborative discussions about student work.

The second paper (Limbrick & Knight, 2005) reported a New Zealand study focussing on “raising student achievement in writing with professional development of teachers viewed as being complementary to the process” (p. 9). The third (L. Reid, 2007) described a Scottish practice of teachers meeting “to negotiate commonly agreed standards for the summative assessment of writing” (p. 132). The teachers reported the benefits of the meetings as helping them to “articulate and share the subject knowledge they already had” (p. 140).

The relevance of these two projects to the current research was considered in terms of the students whose work was considered, the duration of the projects, the methods employed and their findings.

First, both articles (Limbrick & Knight, 2005; L. Reid, 2007) reported on teachers of primary or junior secondary students, rather than the high-stakes environment of the SACE, which was the focus of the current research.

Second, both projects involved teachers collaborating over a length of time, important for teacher change (Evans, 2014; Guskey, 2000; Hipkins & Robertson, 2011; Loucks-Horsley et al., 2003; Wylie & Hodgen, 2010); Limbrick and Knight (2005) described meetings across two years, while Reid (2007) mentioned that the two interviews occurred 9 months apart. Both articles considered the meetings as

communities: a “community of practice” (L. Reid, 2007, p. 133) or a “professional learning community” (Limbrick & Knight, 2005, p. 5). Both cases appear to satisfy the condition of “sustained mutual relationships” (Wenger, 1999, p. 125) that are indicative of a community. Participation in a community of practice has been identified as an important form of teacher professional learning (e.g., Black, 1998; Hipkins & Robertson, 2011; Lave & Wenger, 1991; Loucks-Horsley et al., 2003; Timperley, 2008; Wei et al., 2009; Wenger, 1999). That both projects described the moderators as a community prompted the question of the extent to which SACE Board quality management processes might foster the development of communities of practice among teachers.

The third aspect of the articles that was considered was the methods employed. Limbrick and Knight (2005) used focus groups while Reid (2007) used semi-structured interviews to “encourage teachers to express viewpoints as fully as they chose” (p. 133). Reid also made use of field notes taken during the meetings which were part of the research. The current research employed both methods, semi-structured interviews of teachers and focus groups for the SACE Board officers, as well as field notes made during quality management meetings. The methods employed in these two projects also suggested two ways of determining whether learning had occurred, relevant to the fourth research question in the current research. Reid (2007) used two rounds of interviews, at the beginning and end of the project. The use of more than one round of interviews, with the possibility of before-and-after comparison might help ascertain whether learning had occurred over the research period.

Both papers reported professional benefit to the participants. Limbrick and Knight (2005) described the establishment of shared understandings and greater teacher confidence and knowledge, while Reid (2007) reported enhanced subject knowledge and improved teacher confidence “in their own assessment judgements” (p. 141).

Both projects involved teachers of primary and junior secondary students who were sharing the work of their own students. Teachers referred to the classroom context and the nature of the students whose work was being moderated. Limbrick and Knight (2005) reported competitive, aggressive and non-negotiable teacher behaviour hindering discussions, whereas Reid reported no such problems, possibly

because of “the role of the staff tutor in guiding the discussions” (p. 142). The current research explored a context different in two ways: senior secondary rather than primary and teachers moderating the work of other teachers rather than their own.

The three peer-reviewed papers supported the possibility that teacher learning might occur from involvement in social moderation. The current research explored teachers’ reasons for being involved, or not, in the SACE Board quality management processes and the possible educative value of their experiences.

The literature search had identified gaps which became the subject of the current research. The research reported in the literature differed in three ways from the current research. First, the current research involved the range of quality management processes, not only social moderation. Second, the current research was undertaken in the context of high-stakes senior secondary assessment rather than earlier years of schooling. Third, in the current research, social moderation serves a quality control purpose, rather than quality assurance as in the research reported above.

#### **2.4.2 Other evidence about teacher learning associated with social moderation**

Because the SACE review (Crafter et al., 2006) and others (e.g., Black & Wiliam, 1998; Harlen, 2005) referred to teacher learning associated with social moderation in Queensland, it was considered possible that research had been undertaken in that state on the association between teacher learning and moderation. Consequently, the Australian Digital Theses database was searched for relevant theses. The abstracts of those with titles which appeared promising were read and, where appropriate, further reading was undertaken to select relevant theses.

The search using the terms *professional development* and *teacher* yielded one thesis (Fernandez, 2007). This mentioned two publications arising from it, both in peer-reviewed journals. One of these was dated before thesis completion and the other after. The latter article (Fernandez, Ritchie, & Barker, 2008) was reviewed rather than the thesis. This article focussed on the effectiveness of teacher implementation of a new Physics curriculum in New Zealand, a similar scenario to that of the SACE

changes, although no mention was made of moderation being part of the change. Two points arose. First, Physics teachers were described as belonging to a community of practice, raising the possibility that the current research examine whether the quality management processes might have facilitated SACE teachers learning from community of practice membership. The second point was that the current research, while examining teacher learning, incorporated an element of evaluation of an implementation since it was undertaken in the first year of the new SACE. Fernandez et al. (2008) reported that the implementation under study had been negatively affected by lack of support for teachers and the need to implement a curriculum change “that did not correspond to the shared meanings of many physics teachers” (p. 193). The current research examined the existence of such constraints that might have affected teacher learning.

Searches relating to moderation (social moderation, judgemental moderation, consensus moderation, peer moderation, moderation) yielded another three theses. The earliest (Boereboom, 1999) was set in New Zealand at the time of a movement from national examinations towards “internal standards-based assessment ... [in which] moderation is a key plank” (p. 2). Since the reported research concerned the validity, reliability and manageability of assessment rather than teacher learning, its relevance to the current research was limited. Boereboom’s methods included moderator agreement trials, similar to the verbalisation of assessment practices used in the current research.

The second thesis (Fastier, 2007) reported on semi-structured interviews with teachers during the first two years of the New Zealand National Certificate of Education Achievement (NCEA) for senior secondary assessment. The focus was teachers’ perceptions of the implementation and sought “to identify the factors perceived as supporting or hindering the implementation process” (p. 5). It reported that two years were necessary for teachers to develop effective assessment against the standards, identifying the one-year duration as a limitation of the current research.

The most recent thesis (Barrett, 2008) employed single semi-structured interviews to explore teachers’ collegial activities with teachers of Years 11-13, six years after the innovations of the NCEA, which included social moderation. The thesis reported

that many teachers had commented on moderation contributing to their professional learning.

To find that a number of relevant theses had been written in New Zealand rather than Queensland was, at first, puzzling. On reflection, it might suggest that research is more commonly associated with innovation. Queensland had introduced its SBA system in the 1970s (Maxwell & Cumming, 2011; Wyatt-Smith et al., 2010) and the process has become accepted. On the other hand, New Zealand changed its assessment early this century (Fastier, 2007) and this appears to be associated with the research activity described above. Recent assessment changes in Queensland primary and middle schools (Klenowski & Adie, 2009), in which standards are used to promote consistency in teacher judgement of student work, appear to have stimulated research into social moderation and improvements in teacher judgements (e.g., Adie, 2008; Adie, 2010; Adie, Klenowski, & Wyatt-Smith, 2012; Klenowski & Adie, 2009; Klenowski & Wyatt-Smith, 2010a; Wyatt-Smith et al., 2010). This research is examined in the following section.

The above-mentioned New Zealand and Queensland articles revealed three more sources. The first (Stanley, MacCann, Gardner, Reynolds, & Wild, 2009) reviewed teacher assessment in a number of countries. While reporting professional development arising from teacher involvement in moderation, no evidence was cited to support this claim.

The second article (Davidson, 2003) reported the practice in a Victorian primary school in which, before each reporting period, teachers reviewed the student work of other teachers “to see if they would have drawn conclusions similar to those of the assessing teacher” (p. 83). Teachers provided colleagues with written feedback which could be considered later—no changes were made to student results. Although the article claimed “personal and professional enhancement” (p. 83) as a valuable by-product of the process, the enhancements were not specified.

The third article (Ingvarson, 1990) reported an evaluation of moderation procedures conducted between 1981 and 1984 in Victoria, together with a follow-up study in 1989. The moderation involved senior secondary teachers meeting “in groups of about 10-12 at least three times per year” (p. 6) led by a convenor, with the aim of

“aligning grades with quality of work” (p. 1), and was referred to as *consensus moderation*.<sup>12</sup> Researchers observed moderation meetings and surveyed teachers, principals, and students. Ingvarson (1990) reported that after one year “the balance of evidence in favour of the new procedures was precarious” (p. 8) but that after four years “the situation had changed significantly” (p. 8). More teachers endorsed the capacity of group moderation to enhance teachers’ skills in assessment and teaching, with 76% reporting improved student learning. Although the reported Victorian practices and the SACE practices of 2011 both occurred in high-stakes senior secondary assessment, there were differences. The system reported by Ingvarson (1990) involved all teachers in a series of meetings throughout the year, whereas the new SACE moderation employs a fraction of teachers as moderators, who meet once at the end of the year. The current research investigated the possibility of teacher learning from involvement in all of the quality management processes, not only moderation.

### **2.4.3 Recent evidence for teacher learning from social moderation**

As mentioned above, since the current research began, there have been papers on the association between teacher learning and moderation, arising from developments in New Zealand and Queensland. Although they did not contribute to the framing of the current research, they are relevant and are examined below.

#### **2.4.3.1 *New Zealand research on teacher learning and social moderation***

The New Zealand research was undertaken in the context of the new national curriculum for the NCEA, disseminated in 2006 for implementation by 2010 (Cowie et al., 2009). 2010 also saw the introduction of National Standards for Years 1–8 (Cowie et al., 2009; Hipkins & Hodgen, 2011; Hipkins & Robertson, 2011), although these were not directly linked to the new curriculum (Cowie et al., 2009, p. 49). The papers reporting research on the potential for learning from involvement in social

---

<sup>12</sup> This high-stakes assessment complies with the group moderation previously undertaken by SSABSA (B. Atherton, 2009; Gipps, 1994b), in which teachers participated in quality control discussions of their own students’ work. In the SSABSA version teachers met once each year for approximately three hours, whereas the Victorian teachers met three times each year.

moderation will be examined in chronological order.

The earliest (Hipkins, 2010a) reported on New Zealand Council for Educational Research (NZCER) surveys, which collected data from a range of people about the NCEA. Hipkins (2010a) indicated the importance of teachers' views of moderation, asserting that it is preferable for teachers to perceive moderation as having a learning purpose. Having teachers perceive it as "predominantly for accountability ... squanders the learning opportunity and consolidates the view that incorrectly assigned levels of achievement are mistakes rather than matters of judgement" (p. 31). She explained that all assessment involves matters of judgement that in "traditional examinations ... are likely to be hidden from the view of all except the markers" (p. 13). Respondents reported moderation as time-consuming but generally agreed that it provided valuable insights for teachers. The majority of teachers and principals reported a lack of consistency between moderators, and the paper described the appointment of full-time moderators in response to media comments on the "reliability and fairness of teacher judgements" (p. 31). Secondary teachers reported that recent feedback had been more helpful than previously, as it had clarified the standards required in student work. The report did not describe, however, how this had been achieved.

Later in 2010, Hipkins (2010b) discussed criticisms of the NCEA results, focussing on grading differences between teachers and moderators. Reiterating her comment that similar differences occur between examination markers, she contrasted this with moderation which "brings the whole act of judgement making out into the daylight" (p. 18). She argued that moderation offers greater possibilities for teacher professional learning than examination marking, which was restricted to markers. Describing moderation as a "learning activity" (p. 19), she argued that a standard "can never reside in words on a page" (p. 19), but is comprised of not only the formal definition, but also the range of assessment tasks, student work and the history of judgements built over time.

Wylie and Hodgen (2010), in reporting the NZCER 2010 Primary and Intermediate Schools National Survey, argued that the introduction of National Standards had introduced a "uniformity of expectation" (p. 18) and had "shone a spotlight on how teachers make their judgements" (p. 18). Most primary school teachers reported that

their schools had used some form of moderation prior to the mandated moderation of 2010. Teachers who believed there was a high level of consistency in teacher judgements were more likely to be in schools practising moderation.

Hipkins also led a team that reported (Hipkins, Cowie, Boyd, Keown, & McGee, 2011) on a number of innovative ways in which the revised curriculum had been implemented. Some teachers commented that annotated student work and moderation feedback were “extremely useful” (p. 4) and it was suggested that providing teachers with assessment feedback was one way of building assessment capacity.

Hipkins and Robertson (2011) synthesised a report of projects that explored opportunities for teacher learning when teachers participate in moderation of student work, with the intention that moderation might be as fruitful as possible for teachers. They reiterated that standards are more than written statements; they include examples of tasks and student work as well as the accumulation, over time, of judgements and teacher awareness. Standards emerge gradually. They cited the NCEA experience which “took several years of assessment cycles to consolidate” (p. 6), yet still requires maintenance and fine-tuning. The authors described moderation as a “complex and social process” (p. 16). While acknowledging that social moderation is subjective, they stressed that this does not mean that “anything goes” (p. 16) and argued that its subjective nature needs to be recognised as part of the process rather than undermining its validity and reliability. They argued that “moderation is best structured as an iterative process where shared understandings grow over time” (p. 19), a condition not satisfied by “one-off” meetings. They also noted that moderation “needs to demonstrably lead somewhere constructive ... [which would be more likely] if moderation is focused on improving learning rather than on compliance” (p. 20). Moderators need to see themselves as learners. The paper concluded with the observation that “there are tantalising indications of the potential for powerful professional learning to emerge during moderation conversations, ... that the area of moderation as professional learning is ripe for further research and development” (p. 31). Thus, from New Zealand came support for the purpose of this thesis to investigate the potential for teacher learning in the quality management processes.



Hipkins and Hodgen (2011) analysed responses to the NZCER 2010 national survey on the professional learning outcomes of intra-school moderation, acknowledging that moderation had “largely been under-researched as a professional learning activity” (p. 1). They described moderation as requiring “*collaborative* [italics in original] decisions to reach consensus agreements” (p. 2) and argued that moderation was “an important professional responsibility for *all* [italics in original] New Zealand’s primary school teachers” (p. 2).

The most recent New Zealand article (Smaill, 2013) reviewed the literature with the purpose of determining “the professional learning that is required to ensure the efficacy of ... social moderation” (pp. 250, 251). Thus the focus, in the context of the primary school National Standards, was learning which might be *required for* moderation rather than learning that might be an *outcome* of moderation. From observations of other systems, Smaill noted that “although participation in social moderation can facilitate professional learning, ... it can also result in missed learning opportunities and unintended learning” (p. 251). Smaill sought the causes of these missed opportunities and unintended learning by analysing the social moderation processes in the light of the existing evidence on effective teacher learning. In this she was undertaking a similar task to the second goal of the current research, which was to investigate the SACE Board’s quality management processes in the light of the literature on effective teacher learning.

Smaill argued that social moderation satisfied some of the features associated with effective professional learning opportunities, including “engage in learning over an extended period; use assessment as a catalyst for further learning; challenge prevailing discourses ... and, in some cases, work with external experts” (p. 256). She believed that improvements in teacher assessment could be facilitated by “reconfiguring teacher professional learning as a goal, rather than a by-product, of social moderation” (p. 251).

Smaill contrasted the moderation requirements for “relatively lower stakes” (p. 253) primary schools with those of NCEA (i.e., senior secondary), which include “the submission of assessment materials for pre-moderation, school-based moderation or co-marking of student work” (pp. 252, 253), as well as the submission of marked student work for external moderation. Smaill described the quality assurance

purpose of social moderation as “process oriented” whereas the quality control function was “product oriented” (p. 253).

She indicated that teachers would need to use moderation sessions to review, not only their interpretations and applications of the National Standards (so addressing assessment reliability), but also the validity of their assessment tasks, since common assessment tasks were not prescribed. She reported the benefit of access to external expertise (Timperley et al., 2007) and suggested that such access might assist in ensuring teachers’ conceptual clarity.

Smaill also noted that bias is recognised in teacher judgements and cited research that teachers require training to assist them in identifying potential sources of bias. She argued that, for the potential teacher learning from moderation to be realised, “teachers will require regular opportunities to take part in well facilitated, externally supported, moderation processes” (p. 259).

#### **2.4.3.2 Queensland research on teacher learning and social moderation**

Although Queensland has a long history of externally moderated senior secondary school-based assessment, a series of recent papers described research with teachers of Years 4–9 using performance standards and moderation for the first time, focussing on how teachers make judgements and how the development of shared understandings might be improved, one aspect of teacher learning investigated in the current research. The authors acknowledged an important difference from the senior school practices: the focus was on establishing shared understandings of a standard, rather than the “quality assurance process necessary in high-stakes assessment” (Adie et al., 2012, p. 232). The reforms required teachers to assess student achievement on “centrally-developed assessment tasks” (Klenowski & Wyatt-Smith, 2010a, p. 23), apparently similar to the “controlled assessments” of the UK (Crisp & Green, 2013). This differs from the senior secondary practices in which teachers develop their own assessment tasks (QSA, 2010c). The procedures used for senior schools, apparently for developing teacher skills in assessment task design, were considered “neither logistically nor fiscally possible to replicate” (Wyatt-Smith et al., 2010, p. 62) for the earlier years. In the present research, the quality management processes were examined for evidence that they offered professional development in

task design.

Queensland has emphasised the role of assessment “to promote learning while fulfilling accountability demands” (Connolly et al., 2012, p. 594). Research undertaken in that state might be seen as generating evidence that might contribute to resisting federal challenges (e.g., Ferrari, 2007). It would appear that the SACE Board might also undertake research if it is to defend, or promote, its practices in the national discussion. The current research might contribute to such discussion.

One paper (Klenowski & Wyatt-Smith, 2010a) reported that interpretations of evidence in student work were “influenced by context and the subject domain” (p. 27), reporting that “teachers of English, for example, relied on judgement processes that were significantly different from those of mathematics and science teachers” (p. 33). Mathematics and science teachers used marks in arriving at their judgements, being uncomfortable with standards they regarded as “open to interpretation” (p. 34). The authors observed that mathematics and science teachers were more likely to claim that “objectivity of judgement is realised through numeric scoring” (p. 35). Hipkins and Robertson (2011) reported “differences in the decision-making process for teachers of different subjects” (p. 15) in a Canadian study (Duncan & Noonan, 2007) which reported that mathematics teachers gave less consideration than other teachers to non-cognitive aspects such as student effort and motivation. These observations identified a limitation of the current research discussed later (see 9.3).

Connolly et al. (2012) studied Year 4–9 teachers’ views on the “comparability and consistency of their judgements” (p. 597). Participants agreed that the use of the standards and moderation had achieved consensus, suggesting that learning had occurred as teachers moved towards shared understandings of the performance standards. For the current research this raised the question of what evidence might be used to demonstrate improved consistency between teachers.

Wyatt-Smith et al. (2010) reported that, while teachers made use of the resources provided (e.g., annotated examples, advice on grading), they did so to different extents, and other information (e.g., unstated ‘in the head’ criteria and knowledge of students) was also used in evaluating student work. It was found that, where unstated criteria were used, teachers assumed that the same standards were held by

other teachers. The authors concluded that “the provision of assessment materials does not necessarily lead to common practice or shared understandings” (p. 65). As two of the authors explained in another article, performance standards “are necessary but not sufficient” (Klenowski & Wyatt-Smith, 2010a, p. 21). They suggested that the benefit of performance standards is best realised in association with “opportunities for teachers to develop shared interpretations ..., especially through moderation” (Wyatt-Smith et al., 2010, p. 72). By considering the range of quality management processes, the current research provided an opportunity to investigate the range of resources used by SACE teachers. It also explored whether there was evidence of increased teacher familiarity with some of those resources, specifically the terminology and performance standards of SACE Board documents.

The most recent article (Adie et al., 2012) focussed on whether social moderation led to teachers changing their judgements about student work, exploring whether learning—the development of common interpretations of the performance standards—had occurred through social moderation. The development of common understanding depends on participation in negotiation and the authors described teacher behaviour changing over the research period. In early meetings, teachers relied “heavily on the annotated samples and used them as exemplars of a standard rather than one example of how competence may be demonstrated at that standard” (p. 234). Later they realised that “the annotated samples only provided one possible response and that there were other ways to demonstrate evidence of a standard” (p. 235). The authors concluded that teachers changed their judgements about student work but that the changes were “not always related to ‘the application of explicitly defined standards’” (p. 238). Other factors influencing judgements included the teachers’ own standards, professional knowledge and prior moderation experience. The authors anticipated that, with experience, the reliability and validity of teacher assessments would increase, demonstrating teacher learning. They recognised, however, the need for verification of this outcome. The current research collected data on teachers’ judgements, seeking evidence of improved consistency, which might be related to involvement in the quality management processes.

In summary, Queensland researchers affirmed “it is crucial that guidelines and professional development opportunities be provided to teachers” (Connolly et al.,

2012, p. 613), that moderation is “essential for developing coherence and consistency” (Wyatt-Smith et al., 2010, p. 62) and that moderation could “no longer be considered an optional extra” (Klenowski & Wyatt-Smith, 2010a, p. 21). Their beliefs find expression in the Queensland assessment policy document (QSA, 2009) which states that, for social moderation to work effectively, there is a need for “consensus through teacher discussions on the quality of the assessment instruments and the standards of student work” (p. 3). The current research examined opportunities for such teacher discussions in the range of quality management processes, not only social moderation.

Klenowski and Wyatt-Smith wrote another paper (2010b) that was more theoretical in nature. Rather than reporting research, this examined standards, teacher judgement and moderation within the context of national curriculum and assessment reform. They discussed four conditions to be addressed when standards are implemented, conditions that might have been anticipated in the SACE implementation and so sought in the current research. The first was the need for clarity about the purposes and functions of the standards. The second related to teachers’ understandings; the authors argued that, although “assessments reliant on human decision making are ultimately subjective” (p. 113), they can be made dependable “if standards are promulgated in appropriate forms and teachers have the requisite conceptual tools and professional training” (p. 113). In particular they described the need for exemplars, a topic addressed in more detail later in this chapter. The third condition was that teachers be provided with opportunities to participate in moderation so they can “develop a shared understanding of the meaning of standards” (p. 114). They dismissed the contention that social moderation is an “optional extra” (p. 115) for two reasons: it can provide “the necessary checks and balances” (p. 115) on teachers’ judgements and “it has been shown to build teacher assessment capacity” (p. 115). The fourth condition was the need for teachers to be provided with guidelines and professional development opportunities, both of which might be expected in the SACE Board quality management processes. Thus, Klenowski and Wyatt-Smith (2010b) saw standards as incorporating exemplars, shared experiences and practice accumulated over time.

The Queensland research, begun prior to the current research but reported during it,

is relevant to this thesis, although it reports on work with primary school teachers rather than senior secondary. It reiterated the need for professional development for teachers in the development of assessment tasks and in the use of performance standards. The Queensland research concentrated on the development of common standards through social moderation whereas the current research looked for evidence of teacher learning through participation in a wider range of activities. From the Klenowski and Wyatt-Smith paper (2010a) a limitation in the current research was recognised, as all participants were teachers of the same subject, Physics. The Queensland research also identified conditions which might be looked for in the SACE Board quality management processes: whether teachers were informed about the purposes, advantages and disadvantages of performance standards and whether opportunities were provided for the development of common understanding among teachers.

#### **2.4.4 A summary and analysis of the evidence for teacher learning from quality management processes**

The literature search sought empirical evidence for possible teacher learning associated with the quality management processes. While the search found some evidence linking teacher learning with social moderation, there appears to be a lack of literature on the potential for teacher learning from other quality management processes. The current research addressed this deficit while also contributing to the literature on teacher learning associated with social moderation.

The majority of the literature discussed possible teacher learning from social moderation involving teachers of primary or junior secondary students. Only Barrett (2008) and Ingvarson (1990) dealt with senior secondary teachers. It was also noted that all papers concerned teachers discussing their own students' work. Most discussions served a quality assurance role, with no power to change the grades assigned. Where the grade could be changed, the teacher was involved in the discussion. The current research involved high-stakes end-of-secondary assessment, with teachers having no involvement in their own students' moderation and with grades subject to change.

Teachers in all contexts reported that moderation discussions had contributed to their professional learning. Mention was also made, however, of competitive and

aggressive teacher behaviour hindering discussion, with facilitators serving an important role.

Many authors indicated that standards are more than written descriptors; they include the range of assessment tasks, curricular documents, exemplars of student work and the history of judgements built over time. A number had commented that adequate resources (e.g., time, access to expertise) are essential if the potential educative benefits of social moderation are to be realised.

Research has shown that different teachers use resources differently, with some also making use of their own “in-the-head” standards and knowledge of the students when making judgements. Differences between teachers of different faculties were reported, with mathematics and science teachers more likely to use marks in arriving at judgements, rather than holistic judgements. Many authors mentioned the necessity of discussion in arriving at common interpretations of standards; teachers need to “own” the standards rather than having them imposed.

The attitude teachers bring to moderation was important, with it preferable for teachers to perceive moderation as contributing to learning and not restricted to accountability. Post-moderation feedback served an educative purpose but needed to be carefully constructed for learning to eventuate.

While the review above is not exhaustive, it raises the question of whether it would support Adie’s (2008) contention that “There currently exists little empirical evidence linking teachers’ participation in social moderation processes with the development of common understandings” (p. 5). Has there been enough research to establish that teacher learning is associated with social moderation and to recognise factors that might contribute to this learning? Adie’s (2008) statement that studies had “mainly focused on small groups of teachers and localised practices” (p. 5) is reasonable, although Ingvarson (1990) collected data from 100 teachers. Few studies have involved student work in a high-stakes assessment system. The current research attempted to address this gap and address other quality management processes as well as social moderation.

Having searched for combinations of teacher learning and moderation, subsequent searches were undertaken for these in isolation. Searches were limited to peer-

reviewed articles written after 2000. Using the term *moderation* yielded a few articles which, apart from those listed above, were generally from Queensland authors, with Adie, Klenowski, and Wyatt-Smith being the most prolific. Searches on *professional development* yielded too many articles so the search was restricted by addition of the terms *teacher* and *effective*. Authors such as Borko, Desimone, Garet, Guskey, and Porter led to other articles referred to previously.

The next two sections describe some aspects of quality management of SBA and teacher learning not addressed above.

## **2.5 Quality management of school-based assessment**

The literature search for empirical evidence of teacher learning from quality management processes has been limited to social moderation. The few papers that referred to quality assurance and quality review are now used to examine the SACE Board's processes in these areas.

### **2.5.1 Quality assurance of school-based assessment**

Harlen (1994) identified a number of methods for ensuring that assessment undertaken by teachers is of a satisfactory quality: defining the assessment criteria, meetings of teachers, providing exemplars and visiting advisors. The first two were addressed earlier. The following sections examine the relevance to the current research of exemplars and visiting advisors.

#### **2.5.1.1 Exemplars**

Recommendation of exemplars, serving a number of purposes, to assist teachers in SBA are widespread (e.g., Harlen, 1994; Hipkins & Hodgen, 2011; Klenowski & Wyatt-Smith, 2010a; Sadler, 1987).

Exemplars assist to make explicit the teacher's knowledge of criteria and standards that may exist "in an unarticulated form" (Sadler, 1998, p. 81). Crisp (2010a) argued that they "provide assessors with reference points against which comparisons can be made and hence can help to fix the standard" (p. 53). Sadler (1987) maintained that, for multiple criteria, many exemplars are necessary, a view endorsed by Hipkins and Hodgen's (2011) reference to "a *collective* of exemplars" (p. 2) and Stanley et al.'s (2009) "bank of exemplar materials" (p. 35). Exemplars consist of marked student



work (Harlen, 1994; Hipkins & Robertson, 2011; Klenowski & Wyatt-Smith, 2010b), with annotations explaining how the judgements were made. Klenowski and Wyatt-Smith (2010b) specified that “annotated samples of each standard (A–E)” (p. 119) were necessary, providing “many examples in which the same concept is at work” (Bransford et al., 1999, p. 20).

Exemplars also allow teachers “to know what a change looks like in practice” (Hargreaves et al., 2001, p. 118) so they can apply innovations to their own practices. Hipkins (2010a) described use of exemplars “intended to model innovative assessment practice” (p. 30). Stanley et al. (2009) argued that the exemplars should be “electronically available and regularly augmented and refreshed” (p. 80).

The SACE Board did not mention exemplars in its quality management strategy, nor in documents related to quality assurance (SACE Board of South Australia, 2010f, 2012e, 2012h). Nevertheless, the Board acknowledged a responsibility to provide exemplars (SACE Board of South Australia, 2012f) and has exemplars on its website. Given that the research explored the potential for teacher learning from SACE Board activities, it was decided to include the use of the exemplars, giving eight ways in which teachers might be involved in quality management processes. The current research examined use of the exemplars and the extent to which they satisfied the requirements advocated in the literature.

### **2.5.1.2 Visits by advisors**

Previously, SSABSA employed *support moderators* (SSABSA, 2006) in some subjects to provide “support and guidance to teachers on their assessment standards ... [and] to ensure a common understanding of the assessment criteria through inspection of marked student work, approval of assessment plans, assessment tasks, and – in some subjects – teaching programs” (SSABSA, 2006). In earlier research (B. Atherton, 2009), teachers reported high regard for support moderators but believed that the role was under threat, that schools visits and “the consequent face-to-face discussion had been curtailed” (p. 104).

The teachers were proven right, for the new SACE has no support moderators (*future SACE Office*, 2008). The lack of support moderators was, however, considered unlikely to arise in the current research since Physics was a subject which had never

had support moderators.

### **2.5.2 Quality review of school-based assessment**

*Quality review* refers to processes that occur after assessments are completed and are undertaken for the purpose of improvement. Queensland undertakes “random sampling” (QSA, 2010a), in which a random selection of student folios are independently reviewed by other district panels. The results are published annually, reporting a “high level of consistency of teacher judgments” (QSA, 2010a, p. 10). Given the apparent lack of similar data from other systems, it is not clear how others’ external assessments compare with Queensland’s.

The SACE Board refers to *post hoc* processes which constitute the *improving* phase of its four-phase cycle, in which the data are “monitored and analysed to develop improvement strategies” and schools are provided with data “to enable them to monitor and review each phase of their own quality assurance processes” (SACE Board of South Australia, 2010f).

The SACE Policy Framework (SACE Board of South Australia, 2013e) acknowledges responsibility for “preparing and publishing information about the moderation of school assessment in a form that helps teachers to strengthen the use of performance standards in their teaching” (p. 21) while specifying that teachers have the responsibility of “analysing and using the results ... to strengthen their understanding and use of the performance standards” (p. 22). As mentioned previously, the current research collected data on the moderation feedback and its potential for teacher learning.

This section about quality management has considered its three components—*quality assurance*, *quality control*, and *quality review*—and their potential for teacher learning.

## **2.6 Effective professional learning for teachers**

The quality management processes were examined, in light of the literature on effective teacher learning, to ascertain whether, and if so how, they might have contributed. This section begins by discussing the terms *professional development* and *teacher learning* and defines their use in this thesis. It then presents an analysis

of the literature and the factors found to contribute to effective professional development activities and effective teacher learning, which were then used in the current research.

### **2.6.1 Terminology associated with teacher learning**

Much of the literature about teacher professional learning discusses *professional development*. Some authors, however, (e.g., Desimone, Porter, Garet, Yoon, & Birman, 2002; Garet et al., 1999; Garet et al., 2001), although reporting research on factors contributing to the effectiveness of *professional development*, did not always define their use of the term. Guskey (2000) defined *professional development* as “those processes and activities designed to enhance the professional knowledge, skills, and attitudes of educators” (p. 16). He also, however, described professional development as “an ongoing and continuous process, not an event” (2000, p. 16), suggesting that it is more than designed activities. Fullan and Stiegelbauer (1991) defined professional development as the “sum total of formal and informal learning experiences” (p. 326) throughout a teacher’s career, which appears to align with Guskey’s “ongoing” process rather than his “designed activities”. Fullan, however, in a later edition (2007) of the same book, discriminated between *professional development* and *professional learning*, stressing that “professional learning is not about workshops and courses” (p. 283). In this thesis the term *professional development* will be used to refer to “activities designed to enhance the professional knowledge, skills, and attitudes of educators” (Guskey, 2000, p. 16).

*Professional learning*, however, can occur in many ways, not only through planned professional development activities. Fullan (2007), for example, mentioned that “teachers are often the preferred source of ideas for other teachers” (p. 75) while Timperley (2008) commented that “many studies suggest that participation in a professional community with one’s colleagues is an integral part of professional learning” (p. 19). Pollard (2002) was more specific about what *enhance* might mean: increasing one’s understanding or ability to apply the learnt material. In this thesis, the term *teacher learning* will be used for changes in teachers’ assessment beliefs and practices, since assessment was the focus of the SACE Board’s quality assurance processes.

The next sections describe the features of effective professional development

activities and their relevance to the current research.

### **2.6.2 Effective professional development activities**

Much has been written on the benefits of professional development for teachers (e.g., Darling-Hammond & McLaughlin, 1999; Hawley & Valli, 1999) and there is recognition that “mere attendance [at professional development activities] ... will not necessarily result in enhanced learning” (Díaz-Maggioli, 2004, p. 130). Recent decades have seen research into features associated with the effectiveness of professional development activities. This section analyses a number of research reports that investigated these associations.

Garet et al. (1999) reported a consensus among researchers concerning the features contributing to effective teacher professional development. They described how “traditional approaches to professional development” (p. 1-4)<sup>13</sup>, commonly conferences or short workshops, had failed to induce significant change in teacher practice. Consensus was that, in order to be effective, professional development activities needed to be of sufficient duration and that certain forms were likely to be more effective because they provided prolonged opportunity for teacher discussion. These preferred forms included study groups, coaching, mentoring and networking. They also cited research showing that “professional development focusing on subject matter content and how children learn that content is effective in boosting student achievement” (p. 1-5). Garet et al. (1999) also studied the effect of “collective participation” (p. 3-5) of teachers from the same school, department, or grade level, as opposed to participation of individuals from different schools.

Garet et al. (1999) reported on a cross-sectional study of teachers involved in the Eisenhower Professional Development Program. Teachers were asked to describe the professional development activities in which they had participated. In another paper (Garet et al., 2001) teacher outcomes were described as “increases in knowledge and skills ... and changes in classroom practice” (p. 920).

Garet et al. (2001) described the professional development activities in terms of

---

<sup>13</sup> Garet et al. (1999) numbered the pages in each chapter from 1, for example p. 3-5 is p. 5 of Chapter 3.

“structural features” and “core features” (p. 919). They analysed the data to find those features having significant positive effects on teacher knowledge, skills and classroom practice. Their findings were as follows.

### **2.6.2.1 Structural features**

The structural features of professional development activities consisted of the form of the activity, its duration and collective participation by teachers (Garet et al., 1999). The authors classified the forms of the activities into two categories: traditional forms and those which were more participatory in nature. The traditional form included workshops and conferences tending to take place outside the teacher’s school or classroom, involving leaders with special expertise, and with attendance at scheduled times. Hawley and Valli (1999) described them as “brief, often one-shot sessions” (p. 134). The “reform type” (p. 3-8) of activity included study groups, coaching, mentoring, and networking. Hawley and Valli (1999) described reform activities as “a shared public process” with teachers as “active participants”. They reported reform activities as often occurring during the school day as part of a teacher’s regular work (p. 3-9). They believed that locating professional development opportunities in the teacher’s workplace would more likely lead to teachers making “connections with classroom teaching” (p. 3-9) and might contribute to sustained learning over time.

The duration of the activities referred to the total amount of contact time and the period over which the activity occurred. The contact time (*contact hours* (Garet et al., 1999)) was important because greater time provided increased opportunities for “in-depth discussion” (p. 3-12), while a series of activities over a longer time (*time span* (Garet et al., 1999)) allowed opportunities for teachers “to try out new practices in the classroom and obtain feedback” (p. 3-12). The authors did not define the contact hours or time span sufficient for effective activities, but their overall mean was 25 contact hours (p. 3-13) and 2–4 days was the most common time span (p. 3-15). For the current research it was decided that 25 hours would be the necessary contact hours and 4 days would be the minimum time span for effective professional development activities. The 25 hours is close to the 30 hours identified by Darling-Hammond and Richardson (cited in Wells, 2014) as the minimum time for effective professional learning, but four days is well short of the six months Darling-

Hammond and Richardson reported as the necessary time span.

The third structural feature Garet et al. (1999) studied was *collective participation* (p. 3-15) which reflected the extent to which the activities were designed for teachers from the same school, department or grade level. It was believed that teachers who worked together had more opportunity for discussion relating to “concepts, skills, and problems that arise during their professional development experiences” (p. 3-15), while teachers of the same grade or subject were more likely to “share common curriculum materials, course offerings, and assessment requirements” (p. 3-15). Garet et al. (2001) reported that “the collective participation of groups of teachers from the same school, subject, or grade ... [were indirectly] related to improvements in teacher knowledge and skills and [led to] changes in classroom practice” (p. 936). The quality management processes were examined for the extent to which they catered for collective participation.

#### **2.6.2.2 Core features**

The core features of professional development activities (Garet et al., 1999) were a focus on subject knowledge, opportunities for active learning and coherence with other learning activities.

Focus on subject knowledge referred to the extent to which the activity aimed at “improving and deepening” (p. 3-5) teacher knowledge of the subject and how to teach it. The belief that this would improve teacher knowledge, skills and practice was based on reports that a focus on specific content and how students learn had delivered positive effects on student achievement, especially their conceptual understanding (Garet et al., 2001). Since the quality management processes dealt with assessment, the current research interpreted a subject focus as a focus on subject-specific assessment.

The term *active learning* was used by Garet et al. (1999) for a variety of activities involving “meaningful analysis of teaching and learning” (p. 3-5). Examples included reviewing student work, preparation of lesson plans, discussion of classroom implementation, observation of teaching, leading discussions and engagement in written work. The current research considered active learning to include reviewing or discussing LAPs, assessment tasks and student work.

Garet et al.' (1999) third core feature, *coherence*, referred to incorporating into the professional development activity “experiences that are consistent with teachers’ goals, aligned with state standards and assessments, and encouraging of continuing professional communication among teachers” (p. 3-5). In their research, they asked teachers about the degree to which the activity connected with previous and subsequent activities, the extent to which the activity aligned with relevant syllabuses and the extent to which it encouraged networking. The current research examined the quality management processes for these features.

Another strand of research into the Eisenhower project was a longitudinal study (Porter, Garet, Desimone, Yoon, & Birman, 2000). Findings reinforced the importance of the six features identified in the cross-sectional study (Garet et al., 1999) and showed that effective professional development activities influenced teachers’ classroom practices.

More recently the US National Staff Development Council has begun enquiring into teacher professional learning. The first report (Wei et al., 2009) reviewed research linking teacher learning with student learning. This supported the findings of Garet and associates (Desimone et al., 2002; Garet et al., 1999; Garet et al., 2001) regarding the factors necessary for effective professional development activities.

Professional development was examined in the *Teaching and Learning International Survey* (2009), commonly referred as TALIS, which was organised by the Organisation for Economic Co-Operation and Development (OECD). This reports statistical relationships between factors, with no implication of causality. One relationship mentioned is that teachers report as most effective those activities to which they devote the most time. While this apparently supports Garet et al.’s (1999) findings that duration affects the effectiveness of professional development, possibly teachers choose to spend more time on the activities believed to be more effective or which they enjoy.

Desimone (2009) indicated that the effects of the traditional or reform nature of an activity were explained by other features. She suggested that future research investigates subject focus, active learning, coherence, duration, and collective participation. The current research explored the extent to which the quality

management processes possessed these features.

This section has examined the features contributing to the effectiveness of professional development activities. The next section explores the literature on other forms of teacher professional learning and what can be done to ensure such learning is effective.

### **2.6.3 Effective teacher learning**

Recently there has been an increased interest in forms of teacher learning other than “traditional” professional development activities. Much teacher learning arises from contact with other teachers (Evans, 2014; Fullan, 2007; Timperley, 2008; Webster-Wright, 2009). Hawley and Valli (1999) contrasted the “old” paradigm of one-off workshops with teachers passively listening to an “expert” with the “new”, involving teachers in sustained collaboration in a slow process of change. Wei et al. (2010) described the “new” as “a cycle of continuous improvement” (p. 2).

Timperley et al. (2007) and Wei et al. (2009) reviewed research into effective teacher learning. The former examined studies linking student outcomes to teacher professional learning and development, while the latter included case studies and qualitative methodologies as well as experimental studies. In particular, they reported teachers learning from “professional learning communities” (p. 9), peer observation, analysis of student work and student data, as well as coaching and mentoring.

Desimone (2009) also summarised the features of effective teacher learning found in the “empirical research base” (p. 181). The findings of these three (Desimone, 2009; Timperley et al., 2007; Wei et al., 2009) are summarised in Table 1, with the conclusions of Wei et al. (2009) taken from Wei et al. (2010, pp. 6, 7).



**Table 1: Features of effective teacher learning**

<b>Desimone (2009)</b>	<b>Timperley et al. (2007)</b>	<b>Wei et al. (2009, 2010)</b>
Content focus		Focus on content and pedagogy
Active learning	Focusing on engaging teachers in the learning process	Active learning
Coherence	Content consistent with wider policy	Coherence
Duration	Sufficient time, using time effectively	Intensive, sustained over time
Collective participation		
	Opportunities to interact in a professional community	Professional learning communities
		Embedded in teachers' day-to-day work
		Related directly to teachers' work with students
	Engaging external expertise	
	Challenging problematic discourses	

The previous section examined aspects of the quality management processes identified in Table 1: subject focus, active learning, coherence, duration and collective participation. The following section explores the quality management processes for evidence of the other features identified in Table 1: encouraging membership of a community of practice, relationship to teachers' school practices, access to external expertise and challenging problematic discourses.

Timperley et al. (2007) frequently referred to feedback assisting teachers to translate theory into practice although not including it in the elements identified in Table 1 (Timperley et al., 2007, p. xxvi). The benefit of feedback was also reported by others (Davidson, 2003; Garet et al., 1999; Hipkins, 2010a; Hipkins & Robertson, 2011). Consequently, the use of feedback in the quality management processes was also explored.

### 2.6.3.1 Professional learning communities

Although Mindich and Lieberman (2012) asserted that “Historically, teaching has been an isolated activity” (p. 3), the last twenty years has seen evidence accumulate on the educative benefits of collaboration among teachers.

Timperley et al. (2007) and Wei et al. (2010) reported research demonstrating the educative benefits of an ongoing learning relationship between teachers. Wei et al. (2009) described teachers practising and researching together to “support one another in implementing new ideas” (p. 13) and Timperley et al. (2007) described teachers “working together regularly ... [in a] reform approach to teaching” (p. 74). Starkey et al. (2009) reported that, when assessment system changes were introduced in New Zealand, “teachers spoke positively about the value of meeting with other teachers from their subject area” (p. 184) and that this was most evident among teachers who were the sole subject specialists in their school.

Ball and Cohen (1999) described the potential for teachers learning from involvement in a community of practice enabling them to “cultivate more substantial professional discourse” (p. 17). Fullan (2007) reported the preference of teachers to learn from each other, but recognised that opportunities for teacher interaction were limited. Guskey (2000) described how the culture of an organisation, its “values, beliefs, and norms” (p. 150), has a powerful influence on teacher learning, while Loucks-Horsley et al. (2003) advised that “without a supportive culture, professional learning of teachers has little chance of survival” (p. 91). The SACE Review described Queensland’s moderation as teachers “sharing good practice in pedagogy and assessment” (Crafter et al., 2006, p. 129) and the recommendation of social moderation appeared to be based, at least in part, on the Panel’s faith in this capacity. Hall and Harding (2002) described the formation of communities of practice from discussion around student work and its assessment.

Wenger (1999) described three characteristics of a community of practice: mutual engagement, joint enterprise and a shared repertoire. Loucks-Horsley et al. (2003) contended that the term *community of practice* might be ascribed to “collaborative, school-based communities” (p. 92) although Fernandez et al. (2008) considered New Zealand physics teachers to constitute a community of practice through “the traditions and aspects of the subject discipline” (p. 206). Crisp (2010a) argued that

teachers implementing SBA could be considered a community of practice, although membership might not be as explicit as in examination marking. She suggested they constituted “seasonal, dissipated communities of practice” (p. 44).

Lave and Wenger (1991) used the term “legitimate peripheral participation” (p. 29) to describe the ways in which newcomers become part of a community of practice. This process allows newcomers opportunities to “gradually assemble a general idea of what constitutes the practice of the community” (p. 95). Members “participate in a community of practitioners as well as in productive activity” (p. 110). In the context of the quality management processes the productive activity might comprise discussing the application of performance standards in assessing student work. As others (Fuller, Hodkinson, Hodkinson, & Unwin, 2005) have indicated, Lave and Wenger were working with stable, established communities of experienced practitioners with newcomers being introduced through “legitimate peripheral participation”. In the current research, the situation revealed many of the complexities described by Fuller et al.. First, the participants were already members of communities of practice in their own schools. Second, whether a prior community of practice existed among SACE Physics teachers is questionable. There was no moderation before 2011 and any communities of assessment practice for setting, vetting and marking examinations were limited in membership and duration.<sup>14</sup> Third, many “newcomers” to the quality management processes were highly experienced teachers, differing from the usual concept of a “newcomer.” Thus the role of the teachers was complex. In the teaching of Physics and in previous assessment practices, many would likely consider themselves “expert”, whereas in the new assessment system that uses performance standards and social moderation, Physics teachers were “newcomers” and needed to be inducted into these practices, using their expertise. In this they would comply with Lave and Wenger’s (1991) description of newcomer as they “gradually assemble a general idea of what constitutes the practice of a community” (p. 95). It was considered that the quality management processes might afford opportunities for teachers to meet and exchange ideas in ways not previously available, thus facilitating the development of a

---

<sup>14</sup> Each year three or four setters meet approximately four times in February–March and a different group of three or four vetters have one or two meetings in April. In November approximately 25 markers meet twice.

community of practice among Physics teachers.

Online networking can support learning and reflection (Bransford et al., 1999) and SACE Board online forums have existed in all subjects since before the new SACE. It was possible that the SACE Board and Physics teachers might use the online forum to communicate with each other as members of a community of practice. There was also the possibility that, as an outcome of the quality management processes, teachers might form local communities of practice. Although such communities could not be described as a SACE Board quality management process, they might be facilitated as a consequence of the changes in assessment. Thus membership of communities of practice was examined in the current research, giving nine ways in which teachers might learn from the quality management processes.

### ***2.6.3.2 Learning opportunities embedded in teachers' practices***

Putnam and Borko (2000) contended that the situation within which learning occurs is an integral part of the learning. From this, they argued that the most meaningful locations for teacher learning are in the teachers' own schools, so grounding learning in the teachers' own practices (Postholm, 2012; Webster-Wright, 2009; Wei et al., 2009). The learning requires sustained collaboration that focuses on deepening teachers' practices (Wei et al., 2009). Wei et al. (2010) reported a consensus regarding the benefit of professional learning "embedded in teachers' day-to-day work in schools" (p. 36). Saxe, Gearheart, and Nasir (2001) reported student learning benefits when teachers engaged in "integrated and ongoing professional development" (p. 71), describing the benefit as deepening teacher expertise.

For the current research, the relevant area of practice was SBA. Timperley et al. (2007) reported that approximately half of the research studies reporting positive student outcomes involved teachers "developing their understanding of and use of assessment" (p. xxxiv). The formative use of assessment in these studies, however, differed from the summative assessment that was the focus of the current research. Hipkins and Robertson (2011) argued that discussion of student work would increase teacher awareness of what good student work is like, potentially improving student learning. It was considered that teacher learning in the area of summative assessment might be associated with practices similar to the effective practices identified by Timperley et al., regardless of whether they were designed to improve

student learning.

### **2.6.3.3 Challenging prevailing discourses**

Timperley et al. (2007), identified challenging “problematic discourses” (p. xxvi) as an important element of effective teacher learning. They identified two problematic areas: that some students could not learn as well as others and the most effective way of teaching particular curricula. They described these as problematic because “considerable gaps” (p. 73) existed initially between teacher beliefs and what was promoted in the professional learning. They indicated that teachers need powerful reasons “to engage with new information in sufficient depth to change their practice” (p. xxxviii), that current practice should be regarded as “worthy of debate” (p. 194) and that contemplating such challenges might motivate teachers to change. They argued that the status quo was likely to become further entrenched if there was a failure to challenge problematic beliefs.

Timperley et al. (2007) described how resolution of “a degree of dissonance with teachers’ current positions” (p. 89) was a feature of many of the studies they reviewed. “When new information challenges previously held beliefs and values, dissonance is created” (Timperley et al., 2007, p. 13). Thomson and Zeuli (1999) argued that changes in deeply held beliefs required dissonance to be created between teachers’ beliefs and practices, followed by opportunities for teachers to “think through” (Loucks-Horsley et al., 2003, p. 45) the dissonance. Suggested strategies included discussion and reading (Loucks-Horsley et al., 2003), which offer “iterative cycles of thinking” (Timperley et al., 2007, p. xxvii).

The common thread in these papers is the requirement to challenge beliefs that conflict with new practices. For the current research, this raised the question of whether teachers might consider the use of performance standards to conflict with their existing assessment practices. The quality management processes were examined for evidence that they addressed any dissonance and provided opportunities for its resolution.

### **2.6.3.4 Access to expertise**

Timperley et al. (2007) described the involvement of external experts as “necessary but not sufficient” (p. xxvii) in promoting effective teacher learning. They explained

that expecting teachers to think about their practices in “new ways” (p. xxix), required sources that could bring “new perspectives” (p. xxx). They explained how, without access to external expertise, discussions would be restricted to “the collective understandings in the room” (p. 88). To introduce teachers to the use of performance standards, it was desirable for those leading the professional development to have expertise in the use of performance standards (Bransford et al., 1999; Kuijpers et al., 2010).

Earlier reference was made to concerns over bias in SBA (Black, 1998; Crisp, 2010a; Harlen, 2005; Wyatt-Smith, 1999). Harlen (2004a) advised that teachers should be made aware of sources of bias in their assessments and Gipps (1994a) argued that training could help to minimise bias. Timperley et al. (2007) reported that teachers were “typically unaware of the impact of their prevailing discourses” (p. 167) and argued that challenging the discourses was necessary to develop awareness.

Given the bias in SBA (ARG, 2006; Harlen, 2004a), it might be expected that expert input might be used in professional development, possibly diminishing the potential for bias. The current research examined whether the quality management processes included professional development that might assist teachers in identifying and correcting potential sources of bias.

The literature speaks of a number of roles for external experts: providing new knowledge (e.g., on the use of performance standards and teacher bias) and assisting in group discussion. Consequently, access to external expertise was considered relevant in the scrutiny of the quality management processes.

#### **2.6.3.5 Feedback: learning from moderation**

One purpose of performance standards is the capacity for providing feedback that may assist students in their learning (Rolph & Jordan, 2010; SACE Board of South Australia, 2011d). Feedback is important in learning (Bransford et al., 1999) and is among the most powerful influences on learner achievement (Hattie, 2009; Hattie & Timperley, 2007). Effective feedback is also important in teacher learning (Bransford et al., 1999; Davidson, 2003; Garet et al., 1999; Hargreaves et al., 2001; Hipkins, 2010a; Hipkins & Robertson, 2011; McTighe & Emberger, 2008). But not all feedback is of equal value (Hattie, 2009; Hattie & Timperley, 2007). It is most

effective “when it addresses faulty interpretations, not a total lack of understanding” (Hattie & Timperley, 2007, p. 82) and provides specific information that can be acted on (Fogarty & Pete, 2004; Hattie & Timperley, 2007). For teachers who are not moderators, the only opportunity to learn from moderation is from post-moderation feedback. Consequently the current research examined moderation feedback for evidence that it might contribute to teacher learning.

Feedback to teachers need not be restricted to that provided after the end-of-year moderation; it might be available during the year. McTighe and Emberger (2008) advocated that teacher learning might be achieved through collaboration in designing assessment tasks, evaluating student work and improving assessment practices. Wei et al. (2009) reported that teachers learned from defending their practices, as occurs when teachers are involved in moderation discussions concerning their own students’ assessment tasks. Hipkins and Robertson (2011) argued that moderation discussions might generate “a greater understanding of both assessment and moderation processes” (p. 27), as well as increasing teacher self-confidence in their judgements. They argued that skill development was facilitated by ongoing practice, a view supported by Guskey (2002) who maintained that teachers need to receive “regular feedback” (p. 387) on their efforts. Feedback is most effective when timed so recipients might “revise their thinking as they are working on a unit or project” (Bransford et al., 1999, p. 141). Smaill (2013) argued there was a need for “regular opportunities” (p. 259) for teachers to participate in social moderation processes and Hipkins and Robertson (2011) suggested that differences were unlikely to be resolved without such opportunities. SACE Board moderation occurs late in November. Given the evidence above, the best opportunity for teachers to discuss and reflect on their practices is throughout the school year by means of intra-school or inter-school moderation.

Smaill (2013) mentioned intra-school moderation and the submission of assessment materials for pre-moderation as means of ensuring the integrity of high-stakes SBA prior to external moderation, indicating that such practices were used in Queensland, the New Zealand NCEA and Key Stage 3 in Wales. Hipkins and Robertson (2011) described how consistency of teacher judgements in the NCEA took several years, with “pre-assessment moderation” (p. 10) feedback helping teachers to clarify their

understandings. Moderation is a “learning activity” (Hipkins, 2010b, p. 19) because the conversations between teachers “inform assessment practice” (Hipkins, 2010a, p. 32).

The current research examined the quality management processes for two aspects of feedback: whether they provided or encouraged feedback during the year and whether the nature of the feedback might contribute to teacher learning.

#### **2.6.3.6 Quality management processes as effective teacher learning**

Many consider that much teacher learning arises from engagement with other teachers (Allal, 2013; Fullan, 2007; Hargreaves & Shirley, 2009; OECD, 2009; Postholm, 2012). The current research examined the quality management processes for evidence that they encouraged ongoing discussion in the area of assessment. Evidence sought included encouragement of communities of practice, intra- and inter-school moderation during the year, feedback provision to teachers during the year and access to assessment expertise. The quality management processes were also examined for evidence that they challenged teachers’ prevailing assessment understandings and provided opportunities for resolution of any dissonance.

#### **2.6.4 Adult learning**

Studies into adult learning emerged in the 1920s (Knowles, Holton, & Swanson, 2005; Merriam, 2001). A number of authors (e.g., Donavant, 2009; Fogarty & Pete, 2010; Zemke & Zemke, 1995; Zepeda, Parylo, & Bengtson, 2013) acknowledged the leading role played by Knowles in introducing the concept that the principles of adult learning might differ from those applicable to younger people. Knowles used the term *andragogy* to distinguish adult learning from *pedagogy*, the education of young people.

Beginning with four assumptions regarding adult learners, Knowles added others, having six by 1989 (Knowles et al., 2005), as shown in Table 2.



**Table 2: Knowles's assumptions about adult learners**

1. The need to know. Adults need to know why they need to learn something before undertaking to learn it.
2. The learners' self-concept. Adults have a self-concept of being responsible for their own decisions, for their own lives. They resent and resist situations in which they feel others are imposing their wills on them.
3. The role of the learners' experiences. Adults bring to learning greater and more varied experiences than young people. Any group of adults will be more heterogeneous than is true of a group of youths. Hence, greater emphasis in adult education is placed on individualised teaching and learning strategies.
4. Readiness to learn. Adults become ready to learn those things they need to know and be able to do in order to cope effectively with their real-life situations.
5. Orientation to learning. Adults are life-centred in their orientation to learning. Adults are motivated to learn to the extent that they perceive that learning will help them in their life situations.
6. Motivation. The most potent motivators for adults are internal pressures, e.g., the desire for increased job satisfaction or self-esteem.

Most of these assumptions relate to adults' motivation to learn, while acknowledging the importance of previous experience. This is an important element of constructivism, which describes learning as individuals creating their own understandings based on what they already know and believe (Richardson & Placier, 2001). The assumptions also, by placing the learning in the adults' real-life situations, comply with the concept of experiential learning (Kolb, 1984).

Some questioned whether Knowles's work constituted a theory or merely principles of good teaching (Donavant, 2009; Merriam, 2001). Other critics argued that Knowles ignored the socio-cultural aspect of adult learning (Merriam, 2001) or that the principles are not restricted to adults (Donavant, 2009; Merriam, 2001). Later, Knowles (e.g., Knowles et al., 2005) acknowledged that, at times, a pedagogical approach may be appropriate for adults and an andragogical approach appropriate for non-adults.

Donavant (2009) accepted that andragogy was useful for understanding adult learners but argued it was not a "panacea" (p. 230) capable of resolving all adult learning matters. He maintained that adult motivation required learners being convinced of the need for the learning, reflecting Knowles's first principle (Table 2). Donavant argued that imposed learning is resented and "minimally effective" (p. 228).

Merriam (2001) acknowledged andragogy as one important piece in the “mosaic” (p. 3) of theories, models and principles concerning adult learning. She contended that it was as a “guide to practice that andragogy has had its biggest impact” (p. 8).

Zepeda et al. (2013) used the adult learning theory in their analysis of professional development for principals. Having reviewed a number of theories, they identified five common characteristics of adult learning:

- self-directed
- self-motivated
- problem-centred
- relevant
- goal oriented

Some of these are characteristics of adult *learners* rather than of adult *learning*. It was considered desirable to examine the literature for the features of learning experiences appropriate for adults, which, in the above list, are being *problem centred* and *relevant*. In the current research the SACE Board quality management processes were examined for these features.

Knowles (Knowles et al., 2005, pp. 115, 116) advocated a *process* model for adult learning in which the facilitator differs from a teacher in the pedagogical *content* model. Table 3 shows the elements of this model.

**Table 3: Knowles's process model for adult learning**

1. Preparing the learner, e.g., helping the learners to develop realistic expectations.
2. Establishing a relaxed, respectful, collaborative climate conducive to learning.
3. Creating a mechanism for mutual planning.
4. Diagnosing the needs for learning.
5. Formulating program objectives (i.e., content) that will satisfy these needs.
6. Designing a pattern of learning experiences.
7. Conducting these learning experiences with suitable techniques and materials.
8. Evaluating the learning outcomes and re-diagnosing learning needs.

I contend that the first two elements are appropriate for all learners, not only adults.

Elements 3 to 5 concerning negotiation between facilitators (teachers) and learners emphasise the difference from the pedagogical model.

Loucks-Horsley et al. (2003) advised that teacher professional development should be based on effective adult learning principles and recommended the strategies in Table 4.

**Table 4: Designing teacher professional development using adult learning principles (Loucks-Horsley et al., 2003)**

1. Opportunities for learners to provide input to the content and understand the purpose for learning the content which will be addressed.
2. Time for reflection.
3. Multiple modes of presentations and information processing and opportunity to address real problems or challenges.
4. A respect for the expertise adults bring and activities that encourage all to share their knowledge.
5. Support and feedback from people with expertise.
6. Connections between new concepts, information and current knowledge and experience.
7. A safe environment to try new ideas and approaches.

While differences exist between the two approaches, both have learners contributing to the planning of the learning experience. Thus both address Donavant's (2009) assertion that adult educators have a responsibility to assess learners' needs, using learner input in this assessment.

Diaz-Maggioli (2004) observed that adult learner characteristics are often overlooked when planning teacher professional development. He recommended that planning include aims that are relevant to the learners' needs, congruence with the learners' prior experiences and the use of activities and materials that "actively involve participants" (p. 138). Relevance to learners' needs and congruence with prior experiences would be addressed through involving learners in the planning.

While Knowles et al. (2005) did not appear to advise on learning strategies, similarities exist between "multiple modes of presentations ... and opportunity to address real problems" (Loucks-Horsley et al., 2003, p. 245) and activities and materials that "actively involve participants" (Díaz-Maggioli, 2004, p. 138).

The principles above were endorsed by Zemke and Zemke (1988) who argued that

adult learning should be integrated into the learner's previous knowledge. They recommended active learning through dialogue with peers and acknowledged the need to "clarify and articulate all expectations" (p. 60) at the start of the learning and later added the need to assess early in the program "where individuals stand on critical issues" (Zemke & Zemke, 1995, p. 34), which is particularly relevant to the need identified earlier to challenge prevailing conceptions.

Analysis of the SACE Board quality management processes sought evidence of the following elements of effective adult learning from those above. First was teachers having an opportunity to provide input into the planning of the learning, thus alerting facilitators to teachers' assessment knowledge and their opinions on the changes being introduced into SACE assessment. Two other elements were relevance to teachers' needs and congruence with their prior experience. The fourth element was the use of appropriate activities and materials, which would address real problems and include active learning and discussion.

The literature was also scrutinised for information on the nature of teacher learning. The next section considers this literature and explains how it was used to examine the SACE Board processes in the current research.

### **2.6.5 The quality of teacher learning**

Marton and Saljo (1976) investigated the learning processes students employed when they were provided with an article and then questioned about it. The researchers aimed to explore "qualitative differences in what is learned" (p. 10) rather than how much had been learned, described as the "traditional method" (p. 4) of describing learning outcomes. Their "most important conclusion ... [was] that learning should be described in terms of its content" (p. 10). They found differences in the quality of learning to be associated with different *levels of processing: deep-level and surface-level*. Later, while retaining the terms *deep* and *surface*, they referred to them as *approaches to learning* (2005). Entwistle (1998) credited Biggs and Ramsden with adding a third category, a *strategic* approach, although Entwistle referred to this as an approach to *studying* rather than to *learning*. In the *strategic* approach a student relies on organisation and an "alertness to assessment requirements" (Entwistle, 2000, p. 2) to achieve the highest possible grades.

Marton and Saljo (1976) described the learning outcomes as four categories concerning comprehension of the *intent* of the article, rather than its *content*. Entwistle (2000) described the surface approach as memorising “unrelated bits of information” (p. 3) whereas the deep approach seeks to relate ideas and uses evidence to examine the logic of an argument. Moon (2000) described the surface approach as involving rote learning and routine use of procedures with no goal of making overall sense of new information, whereas the deep approach involves looking for meaning and constructing personal knowledge, with the new knowledge integrated with previous knowledge. Hattie (2003) described the difference as “surface learning is more about the content ... and deep learning more about understanding” (p. 9).

Although the term *deep* was used by Marton and Saljo (1976) to refer to an *approach* to learning, its use has evolved and it is also currently employed to describe the *quality* of the learning. Thus, for example, Entwistle, McCune and Hounsell (2002), while referring to “deep approaches to learning” (p. 2), also use the terms “deep learning” (p. 2) and “deep level of understanding” (p. 3). Marton and Saljo (2005) also wrote of the “‘depth’ dimension ... implicit in the hierarchies of learning outcomes” (p. 45).

The current research asked two questions about the quality management processes regarding deep learning. First, what evidence was there that activities and materials employed in the quality management processes might facilitate deep learning? Second, what evidence was there of teacher learning of a deep nature rather than of a surface or strategic nature? The next two sections examine what evidence could address these questions.

#### **2.6.5.1 Encouraging deep learning**

Some teaching strategies found to be related to students engaging in a deep approach to their learning might be anticipated in the quality management processes. Other strategies conducive to deep learning were considered not relevant to the current research, including forms of assessment of the learning (Entwistle, 1988), integration of content across disciplines (Ryan, Irwin, Bannon, Mulholland, & Baird, 2004) and work load (Entwistle et al., 2002).

Research on teaching encouraging a deep approach appears to be mainly focussed on tertiary education (e.g., Biggs, Tang, & Society for Research into Higher Education, 2011; Ryan et al., 2004). Biggs and Tang (2011) advised that deep learning teaching required an active response from students, building on previous knowledge and confronting and eradicating learner misconceptions. Ryan et al., (2004) found that references to further reading and discussions with lecturers after class had strong positive associations with a deep approach. Biggs (1999) argued that one way of encouraging a deep approach to learning was to have students “work collaboratively and in dialogue with others, both peers and teachers” (p. 61). A consolidation of writings about encouraging deep learning (University of Technology Sydney, 2012) recommended active engagement, making connections with or challenging prior knowledge and providing opportunities for discussion with other students and staff.

Predictably, some of the strategies encouraging a deep approach are consistent with features of effective professional development activities and adult learning. For example, Garet et al. (1999, 2001) (effective professional development activities), Biggs and Tang (2011) (deep learning) and Diaz-Maggioli (2004) (adult learning) all referred to the active involvement of learners. Although the terms used varied between authors, from this point the term *active learning* will be used in this thesis.

Diaz-Maggioli (2004) described the necessity of “congruence with the learner’s prior experience and knowledge” (p. 138) for effective adult learning. Biggs and Tang (2011) advised that “building on what students already know” (p. 27) can encourage deep learning. In this thesis these are considered synonymous.

The current research examined whether the quality management processes possessed those features, discussed above, that might encourage deep learning.

#### **2.6.5.2 Evidence of deep learning**

Marton and Saljo (1976) described their four categories of learning outcomes as “hierarchically related” (p. 8) in terms of student understanding of the author’s intent. Some authors have suggested alternative ways of describing different qualities of learning. Bloom et al. (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956) proposed a taxonomy of knowledge, comprehension, application, analysis, synthesis and evaluation, which was later revised (L. W. Anderson & Bloom, 2001).

It is, however, mostly used in the writing of questions rather than evaluation of learning outcomes (Biggs & Collis, 1982). The SOLO taxonomy (Structure of the Observed Learning Outcome) of Biggs and Collis (1982) categorised knowledge on the basis of the learner making connections and appreciating their significance (J. S. Atherton, 2013). Shulman (2002) asserted that learners possessing more than knowledge and understanding demonstrated capabilities of performance, reflection and judgement.

Thus various authors have suggested that higher-order mental processes might be demonstrated in diverse ways: awareness of an author's intent (Marton and Saljo, 1976), awareness of interconnectivity between information elements (Biggs & Collis, 1982) or the application, analysis and evaluation of information (Bloom et al., 1956; Shulman, 2002).

The data collected in the current research was not considered susceptible to analysis using the SOLO taxonomy (Biggs & Collis, 1982) or Shulman's (2002) table of learning. Rather, it seemed that Moon's (2000) conceptualisation offered a means of judging the quality of teacher learning. Moon (2000) described *deep learning* as comprising *making meaning*, *working with meaning* and *transformative learning* (p. 138). Deep learning involves learners "relating ideas to previous knowledge and experience" (p. 123), as well as critically examining information and conclusions and searching for underlying principles and connections. Thus deep learning is more about understanding than content (Hattie, 2003). Moon (2000) contrasted deep learning to a *surface* approach which, in the context of the quality management processes, would involve routinely following procedures "without reflecting on either purpose or strategy" (p. 122) or a *strategic* approach with the intention of achieving the highest grades possible. Behaviours associated with a strategic approach included strategies such as completing written work with the "perceived preferences of lecturers" (p. 122) in mind. In the context of the current research this would involve ensuring that student work complied with the perceived preferences of moderators.

Moon (2000) considered that *making meaning* involves the integration of new knowledge into previous knowledge, with linked concepts and "a holistic view of the

subject matter” (p. 143).<sup>15</sup> *Working with meaning*, “manipulation of meaningful knowledge towards a specified end” (p. 144), involves reflection on the new learning, “not necessarily in the physical presence of the new learning” (p. 139). The final stage, *transformative learning*, involves evaluation by the learners of “their frames of references [and] the nature of their own and others’ knowledge” (p. 146). Thomson and Zeuli (1999) defined transformative learning as “changes in deeply held beliefs, knowledge, and habits of practice” (p. 342). Webster-Wright (2009) wrote of the value of critical reflection and “the possibility of transformative change ... through questioning of assumptions that underlie habitual patterns of thought and action” (p. 722). Because the current research was restricted to twelve months, it was considered that changes in deeply held beliefs would be unlikely. Consequently, evidence of transformative learning was restricted to evidence of teachers questioning their assumptions, beliefs and practices.

The current research sought evidence that teacher learning in the area of assessment had involved deep learning (i.e., making meaning, working with meaning and questioning of assumptions, beliefs and practices), which might be associated with involvement in the quality management processes.

### **2.6.6 Teacher characteristics associated with effective teacher learning**

The previously mentioned (see 2.2) interconnected model of teacher learning (D. J. Clarke & Hollingsworth, 2002) described changes in teachers’ beliefs and practices as proceeding through enactment and reflection and referred to an *external domain* offering information, stimulus or support that might contribute to the learning (see Figure 5). The model also described teacher learning without involvement of the external domain, when enactment and reflection occur between the teacher’s knowledge, beliefs and attitudes and professional experimentation. In the current research it was considered that professional growth might be associated with characteristics of the teachers.

---

<sup>15</sup> Terms used by different authors vary. The “active attempt to bring ... past ... knowledge and beliefs to bear in the construction of meaning from present stimuli” has also been called *sense-making* (Spillane, Reiser, & Reimer, 2002) although Moon (2000) used *making sense* to describe gaining a “coherent view of the material in relation to itself, not in relation to previous knowledge” (p. 142). Moon’s *making meaning* and *making sense* will be used in this thesis



Some forms of professional development are more beneficial than others (Desimone et al., 2002; Garet et al., 1999; Hipkins et al., 2011; Wei et al., 2009), representing variation in the external domain of the interconnected model (D. J. Clarke & Hollingsworth, 2002). In the current research, the external domain is provided by the quality management processes and previous sections have described their examination in the light of the literature on effective professional development activities, effective teacher learning and adult learning.

Other researchers have examined teacher factors that contribute to the effectiveness of the learning (e.g., Guskey, 2002; van den Berg, 2002). Many teacher characteristics—such as a teacher’s professional training, the length of teaching experience and the extent of school support—have been investigated for association with participation in professional development and learning from the involvement. For convenience the current research arranged such characteristics into three groups: *professional*, *school* and *personal* characteristics. This section discusses these categories and possible links with teacher learning associated with involvement in the quality management processes.

### **2.6.6.1 Professional characteristics**

An individual’s learning is shaped by the learner’s current knowledge, which may support or hinder the new learning (Bransford et al., 1999). For constructivists, learning involves individuals creating their own understandings based on what they already know and believe, as well as the ideas with which they come into contact (Richardson & Placier, 2001). Thus learners’ prior experiences will influence their involvement in learning opportunities, as well as the degree to which they learn from those opportunities. For teachers, prior teaching experience is one factor influencing teacher willingness to learn. Rosenholtz and Simpson (1990) found that teacher commitment changed “modestly” (p. 252), declining after five years in their careers, before increasing later.

Adult learning is optimised by organising the experience around participant prior knowledge and experience (Díaz-Maggioli, 2004). This suggested that, given background differences, teachers participating in SACE Board quality management processes might report different experiences. Responses to the requirement to use performance standards might vary between experienced and inexperienced teachers.

An experienced teacher might see the performance standards as a hindrance or an affront, while the inexperienced could see them as supportive. Consequently there might be variation in willingness to participate in the processes and openness to any associated learning.

Earlier research (B. Atherton, 2009) reported that teachers' beliefs about moderation were not associated with length of teaching experience. Nevertheless, this was included in the current research, which explored possible associations between teachers' professional backgrounds and involvement in the quality management processes and associated learning. Data were collected on participants' Physics background, teaching experience, Physics teaching experience and prior participation in SACE Board assessment panels.

### **2.6.6.2 School characteristics**

Rosenholtz (1989), in studying "effective schools" (p. 1), collected data on school organisation (e.g., teacher collaboration) as well as teacher background (e.g., years of teaching experience). She found that teachers in "learning-enriched" (p. 98) schools were more likely to spend time and access resources than teachers in "learning-impooverished" schools. She found that 79% of variance in teachers' "learning opportunities were explained by differences in school environment" (p. 102).

Previous research (B. Atherton, 2009) reported that teachers' beliefs about moderation were not associated with teachers' school sectors or locations. Although Greatorex and Suto (2006) found no evidence of significant differences in the marking strategies of teachers with different experiences (p. 21), they cited research that, with greater experience, markers are more likely to use an intuitive approach than a rule-governed process (pp. 5, 21). It was considered possible that similar variations might be found in teachers' practices relating to SBA.

Kilvert and Mercurio (2007), two executive officers of the SACE Board, analysed teachers' participation as SACE Board assessors on the basis of teachers' schools: location (whether the school was metropolitan or regional), sector (whether the school belonged to the Catholic, Government, or Independent sector) and socio-economic data. They reported that, in 2005, 11% of Year 12 teachers were involved in "moderation and external assessment processes" (p. 7) and that differences existed

between the sectors. Government schools were under-represented and metropolitan schools over-represented. To explore further the work of Kilvert and Mercurio (2007), in the current research, data were collected on teacher location and schooling sector.

A school's culture can have a powerful influence on teacher learning (Guskey, 2000; Loucks-Horsley et al., 2003; Rosenholtz, 1989). Rosenholtz (1989) described schools as being of two types: *learning-enriched* and *learning-impooverished*. In the former, teachers viewed their learning as ongoing, whereas in the latter, teaching was seen as "a gift and that it cannot be taught" (Richardson & Placier, 2001, p. 926). In learning-enriched schools, teachers and principals tried to help colleagues, while in learning-impooverished schools new teachers were left to "figure teaching out" (Richardson & Placier, 2001, p. 926). The OECD (2009) also reported close associations between school climate, co-operation between teachers and professional development (p. 6).

Because the effectiveness of teacher learning is strongly linked to the culture of the teacher's school, participants in the current research were asked about factors Rosenholtz (1989) used to identify schools as learning-enriched or learning-impooverished, such as encouragement of professional learning. It was anticipated that, for example, teachers in learning-enriched schools might be more willing to be involved in the quality management processes and a supportive school might contribute to being more amenable to change. In the current research these data were titled *school support*.

In the current research school data collected included participants' school location, school sector and school support. These data were considered relevant in interviewee selection and in examining possible association with involvement in, and learning from, the quality management processes.

### **2.6.6.3 Personal characteristics**

Personal characteristics have been identified as important in teachers' professional learning (Day & Gu, 2007; Guskey, 2000; van den Berg, 2002). The OECD (2009) acknowledged that teacher participation in professional development activities may be due to teachers' beliefs and found gender differences in the demand for

professional development, in teaching style and in collaborative practices.

Personal data collected in the current research were restricted to gender and a snapshot of participants' assessment beliefs and practices.

#### **2.6.6.4 Factors associated with effective teacher learning**

The data collected on teachers' professional, school and personal characteristics were used to explore possible associations between these characteristics and (1) involvement in the SACE Board's quality management processes and (2) changes in teachers' assessment beliefs and practices.

## **2.7 Summary of research questions**

This chapter began by identifying the four areas of interest in the current research: teacher involvement in the quality management processes (involvement), the form of the quality management processes (form), the content of the quality management processes (content) and teacher learning that might be associated with the quality management processes (learning). The areas evolved during examination of the literature, generating the following four research questions with their sub-questions.

1. What is the extent of teacher involvement in the SACE Board's quality management processes and what teacher characteristics are associated with this involvement?
  - To what extent are teachers involved in the SACE Board quality management processes of the 2011 assessment cycle?
  - What evidence is there of discernible change in teacher involvement in SACE Board assessment panels for the 2011 assessment cycle compared to that of the previous year?
  - What reasons do teachers give for being involved, or not, in the SACE Board quality management processes?
  - What professional characteristics of teachers (i.e., Physics background, teaching experience, Physics teaching experience, prior participation in SACE Board assessment panels) are associated with teacher involvement in the SACE Board's quality management processes?

- What school characteristics (i.e., teacher location, schooling sector, school support) are associated with teacher involvement in the SACE Board's quality management processes?
  - What personal characteristics (i.e., gender, assessment beliefs and practices) are associated with teacher involvement in the SACE Board's quality management processes?
  - What evidence is there of teacher involvement in communities of practice which might be attributed to the quality management processes?
  - To what extent do teachers make use of the SACE Board exemplars and acknowledge their benefit?
2. What evidence is there that the SACE Board's quality management processes possess the form or features of effective professional development identified in the literature?
- To what extent are the quality management processes of a form found to contribute to effective professional development activities for teachers (i.e., active learning, coherence, duration and collective participation)?
  - What evidence is there that the quality management processes incorporate features found to contribute to effective teacher learning (i.e., encouraging ongoing discussion in the area of assessment, encouraging intra- and inter-school moderation during the year, providing feedback to teachers during the year, providing access to assessment expertise and challenging prevailing discourses)?
  - What evidence is there that the form of the quality management processes might contribute to effective adult learning (i.e., consultation in the planning process, relevance to teachers' needs, congruence with prior experience and appropriate activities and materials)?
  - What evidence is there that the quality management processes are of a form that might encourage deep learning (i.e., active learning, building on teacher prior knowledge, confronting and eradicating teacher misconceptions, providing opportunities for discussion with other teachers and providing opportunities for discussions with presenters)?

- To what extent do the SACE Board exemplars satisfy the requirements advocated in the literature as desirable (i.e., marked student work with explanation of judgements, tasks across a range of standards regularly expanded and updated)?
3. What evidence is there that the content of the SACE Board's quality management processes provides opportunities for teacher learning in the area of assessment?
- To what extent do the quality management processes focus on assessment in a particular subject?
  - What content in the quality management processes might inform teachers of the reasons for the changes in assessment?
  - To what extent do the quality management processes provide information about the purposes of performance standards and of the advantages and disadvantages of their use in student assessment?
  - What training in the use of performance standards in task design and in the grading of student work is provided by the SACE Board quality management processes?
  - What training is provided in the quality management processes that might assist teachers in identifying and correcting potential sources of bias?
4. What evidence is there of teacher learning that might be associated with involvement in the SACE Board's quality management processes?
- How do teachers describe the learning that they have undergone in the SACE Board quality management processes?
  - What constraints and affordances can be identified that might have hindered or facilitated teachers learning from involvement in the quality management processes?
  - What evidence is there of teacher awareness of the reasons for the changes made in SACE assessment practices in 2011 that might be attributed to the quality management processes?
  - To what extent are teachers aware of the purposes, advantages and disadvantages of performance standards in designing assessment tasks and in the grading of student work?

- What evidence is there of teachers' increased familiarity with SACE Board terminology and performance standards that might be associated with involvement in the SACE Board quality management processes?
- What evidence is there of changes in teachers' attitudes, beliefs and practices in the area of assessment that might be attributed to the quality management processes?
- What evidence is there of teachers adopting active or passive roles in their learning in the areas of assessment and what factors might have contributed to the roles adopted?
- What evidence is there of increased consistency in teachers' applications of performance standards that might be associated with involvement in the SACE Board quality management processes?
- How do teachers describe the feedback provided after moderation and its potential to contribute to teacher learning in the area of assessment?
- What evidence is there that teacher learning in the area of assessment involves deep learning (i.e., making meaning, working with meaning, and questioning previous assumptions, beliefs and practices), which might be associated with involvement in the quality management processes?
- What evidence is there of changes in teachers' attitudes, beliefs or practices in the area of assessment that might be associated with professional characteristics of teachers (i.e., Physics background, teaching experience, Physics teaching experience, prior participation in SACE Board assessment panels)?
- What evidence is there of changes in teachers' attitudes, beliefs or practices in the area of assessment that might be associated with their schools' characteristics (i.e., teacher location, schooling sector, school support)?
- What evidence is there of changes in teachers' attitudes, beliefs or practices in the area of assessment that might be associated with personal teacher characteristics (i.e., gender, assessment beliefs and practices)?

The next chapter describes the methods used to collect the data to provide evidence-based responses in relation to these questions.

## CHAPTER 3: METHODOLOGY AND METHODS

The research used an integrated mixed methods design with a preponderance of qualitative methods. After describing the research purpose, this chapter outlines the reasons for the selected design, the methods used and the sources accessed. The rationale for participant selection is presented, followed by an explanation of how the methods addressed the research questions. The section continues with the limitations and delimitations of the research and a section on ethical matters. Variation of the conducted research from the original design is then explained and the chapter concludes with explanations of data analysis.

### 3.1 Research purpose

The primary purpose of the research was to investigate the potential for teacher learning in the quality management processes associated with school-based assessment (SBA). The research was undertaken in the context of the South Australian Certificate of Education (SACE). The previous chapter explained how nine quality management processes were examined for their potential to contribute to teacher learning:

- attending a planning forum
- approving schools' Learning and Assessment Plans (LAPs)
- receiving LAP feedback
- attending a clarifying forum
- leading a clarifying forum
- moderating
- receiving moderation feedback
- using website exemplars
- membership of a community of practice.

The potential for teacher learning in these was examined through four research questions.



1. What is the extent of teacher involvement in the SACE Board's quality management processes and what teacher characteristics are associated with this involvement?
2. What evidence is there that the SACE Board's quality management processes possess the form or features of effective professional development identified in the literature?
3. What evidence is there that the content of the SACE Board's quality management processes provides opportunities for teacher learning in the area of assessment?
4. What evidence is there of teacher learning that might be associated with involvement in the SACE Board's quality management processes?

Thus evidence was required on teacher involvement, the form and content of the quality management processes as well as evidence of teacher learning.

### **3.2 Methodology**

The exploratory nature of the research and the variety in the research questions required multiple data sources, sources providing “breadth and depth of understanding and corroboration” (R. B. Johnson, Onwuegbuzie, & Turner, 2007, p. 123). Four available sources—SACE Board documents, observation of the processes, teachers involved in the processes and SACE Board officers—were used in the research.

The range of sources suggested using a number of methods, thus providing a more complete account of the topic (Grammatikopoulos, Zachopoulou, Tsangaridou, Liukkonen, & Pickup, 2008; Greene, Benjamin, & Goodyear, 2001). Findings from multiple methods also “enable researchers to conclude whether or not an aspect of a phenomenon has been accurately measured” (Grammatikopoulos et al., 2008, p. 6). Creswell (2008) claimed that using both quantitative and qualitative methods “provides a better understanding of the research problem and questions than either method by itself” (p. 552).

The decision to use questionnaires allowed data collection from a large teacher sample, “facilitating comparison and statistical aggregation of the data” (Patton,

2002, p. 14). Self-report data, such as questionnaires, have been criticised as “likely to be distorted or incomplete to an unknown degree” (J. P. Gall, Gall, & Borg, 2005, p. 180) because respondents may choose to conceal information or may not have the self-awareness to give accurate information. Interviewing a sample of teachers allowed exploration of questionnaire responses, adding depth to the teacher data. Thus, teachers’ data included quantitative (questionnaires) and qualitative (interviews) methods. The other methods (document analysis, observations and focus groups) were qualitative. Consequently, the research used a mixed methods design (R. B. Johnson et al., 2007) with a preponderance of qualitative methods.

Given my extensive experience with teachers as a subject coordinator, a Chief Assessor, an Assessment Field Officer, and earlier research (B. Atherton, 2009), it seemed possible “to be a good interviewer or observer, and learn to make sense of the resulting data, without first engaging in deep epistemological reflection and philosophical study” (Patton, 2002, p. 69). It is possible to separate methods from “the epistemology out of which they have emerged” (Patton, 2002, p. 136). Research that mixes methods in ways “that offer the best opportunities for answering important research questions” (R. B. Johnson & Onwuegbuzie, 2004, p. 18) has been titled *pragmatism* (Creswell, 2003; R. B. Johnson & Onwuegbuzie, 2004; Patton, 2002; Teddlie & Tashakkori, 2009), described as preferring action to philosophising (R. B. Johnson & Onwuegbuzie, 2004). Pragmatism, commonly linked to mixed methods research, focusses on the questions asked rather than the methods and is oriented towards “what works” (Creswell & Plano Clark, 2011, p. 41). It has been argued that methods should be chosen “that are most likely to provide evidence useful for answering important research questions given the inquiry objectives, research context, and the available resources” (Jang, McDougall, Pollon, Herbert, & Russell, 2008, p. 222).

Greene et al. (2001) contrasted pragmatism with a “dialectical stance [which] holds that differences between paradigms exist and are important” (p. 28). Patton (2002) acknowledged that paradigm differences are real, but argued that “paradigm descriptions over-emphasize dualisms such as objectivity-subjectivity” (as reported in Greene et al., 2001, p. 28). It has been argued that the differences between qualitative and quantitative methods are important: “the two components cannot be

logically combined within one single, coherent, and consistent research design” (Bergman, 2011b, p. 101).

My position is that “different methods are appropriate for different situations” (Patton, 2002, p. 72). The use of a variety of appropriate methods is supported by many (e.g., Creswell & Plano Clark, 2011; Greene et al., 2001; Jang et al., 2008; R. B. Johnson & Onwuegbuzie, 2004). Donmoyer (2012a) indicated that Lincoln and Guba had argued that it was “paradigms, not methods, per se, [that] were incommensurable” (2012a, p. 652). While this might be interpreted as falling short of an endorsement of mixed methods research, it implies tolerance.

Creswell (2009) made many points about pragmatism which underlie the approach taken to the present research. The current research was “not committed to any one system of philosophy and reality” (p. 10), giving freedom to choose the most appropriate methods “to provide the best understanding of a research problem” (p. 11), while acknowledging that research “always occurs in social, historical, political, and other contexts” (p. 11). I do not subscribe to either of the purist extremes described by Bergman (2010, 2011a, 2011b). I believe that there is “an external world independent of the mind” (Creswell, 2009, p. 11) but that causal relationships are frequently “transitory and hard to identify” (Teddlie & Tashakkori, 2009, p. 93). As Creswell (2009) explained, pragmatists are “not committed to any one system of philosophy and reality” (p. 10). Instead of asking questions about “reality and the laws of nature” (p. 11, citing Cherryholmes), they prefer to “change the subject” (p. 11, citing Rorty). Creswell summarised pragmatism: it “opens the door to multiple methods, different worldviews, and different assumptions as well as different forms of data collection and analysis” (p. 11).

Possibly because of my science background, I was predisposed towards quantitative research. However, while “quantitative purists maintain that social science inquiry should be objective ... [and that] researchers should eliminate their biases” (R. B. Johnson & Onwuegbuzie, 2004, p. 14), I believe that complete objectivity and elimination of personal bias is impossible in social research: “the conduct of fully objective and value-free research is a myth” (R. B. Johnson & Onwuegbuzie, 2004, p. 16). I do not subscribe to qualitative purists’ contention of “multiple-constructed realities” (R. B. Johnson & Onwuegbuzie, 2004, p. 14) but believe that there is one

reality and different people offer “multiple perspectives” (Phillips & Burbules cited in R. B. Johnson & Onwuegbuzie, 2004, p. 14) rather than multiple realities. I therefore subscribe to mixed methods research, “the third research paradigm” (R. B. Johnson & Onwuegbuzie, 2004, p. 14). I acknowledge that my values may have played “a large role in interpretation of results” (Teddlie & Tashakkori, 2009, p. 8). “All research is interpretive; it is guided by the researcher’s set of beliefs and feelings about the world and how it should be understood and studied” (Denzin & Lincoln, 2005, p. 22). Although “Objective reality can never be captured” (Denzin & Lincoln, 2005, p. 5), the multiple sources and methods were employed in attempting to secure a thorough understanding of the topic (Denzin & Lincoln, 2005).

Using multiple sources, multiple theories of teacher learning and multiple methods of data collection involved triangulation, since it sought “convergence, corroboration, and correspondence of results from different methods” (Creswell & Plano Clark, 2011, p. 62). Although Schwartz-Shea and Yanow (2012) preferred *intertextuality* for analysis “across evidentiary sources” (p. 88), the more common term is *triangulation*. The current research adopted three methods of triangulation: data triangulation (use of a multiple sources), theory triangulation (use of multiple perspectives and theories) and methodological triangulation (use of multiple methods) (Denzin cited in R. B. Johnson et al., 2007). Although Greene, Caracelli and Graham (1989) limited triangulation to instances involving “multiple methods, with offsetting or counteracting biases” (p. 256), I make no claim about “offsetting or counteracting biases”; my approach was more pragmatic—a belief that multiple sources and multiple methods are more likely to give a more complete picture of a phenomenon. For example, after analysing teachers’ comments about the planning forums, new codes were necessary for SACE Board officers’ comments. The officers introduced aspects beyond the knowledge of the teachers, generating a more complete account than could be obtained from either teachers or officers.

Greene et al. (2001) described a number of different ways in which methods may be mixed in research. Because the research questions were written knowing that multiple methods were to be used, each research question could be addressed by multiple methods. During data collection, questionnaire responses, observations,

readings of documents and focus groups were used to develop questions asked in teacher interviews and focus groups; comments made by teachers in interviews were used to generate discussion in focus groups of SACE Board officers. Data analysis revealed facets of teacher involvement that led to the need to undertake new directions not initially envisaged, so that research questions were modified during data collection. Since interaction between the different methods occurred throughout the research, the design was *integrated* (Greene et al., 2001, p. 31), with a dominance of qualitative methods.

The research employed multiple methods of data collection for many quality management processes. Table 5 shows the relationships between the methods and the quality management processes.

**Table 5: Methods used for each of the quality management processes**

		Methods				
		Document analysis	Observation	Teacher questionnaire	Teacher interview	Focus group
Quality management	Planning forum	✓	✓	✓	✓	✓
	LAP approval	✓	✓	✓	✓	✓
	LAP feedback	✓		✓	✓	✓
	Leading a clarifying forum	✓	✓	✓	✓	✓
	Attending a clarifying forum	✓	✓	✓	✓	✓
	Moderation	✓		✓	✓	✓
	Receiving moderation feedback	✓		✓	✓	✓

The research examined the quality management processes of 2011, the first year of the new quality management processes, scheduled as shown in Table 6.

**Table 6: Schedule of SACE Board quality management processes**

<b>Date</b>	<b>QM Process</b>
2010 March - April	Planning forums
2010 December 2011 January	LAP approval LAP feedback
2011 April - July	Leading clarifying forums Attending clarifying forums
2011 December	Moderation
2012 February	Moderation feedback

The research included a longitudinal study with three rounds of teacher questionnaires and interviews, and two SACE Board officer focus groups. An important decision concerned the timing of the teacher questionnaires and interviews. Guskey (2000) described five levels of evaluation of teachers' experiences of professional development activities—from immediate responses to improvements in student learning. Immediate responses entail exit surveys while intermediate levels evaluate participant learning, organisation support and use of the new knowledge. Level 4 information, teacher use of new knowledge, was considered the most appropriate level for the current research which investigated the potential for teacher learning in the quality management processes. Guskey recommended that evidence be collected “after sufficient time has passed to allow participants to adapt the new ideas and practices to their setting” (p. 85). Although planning forums were conducted in the first half of 2010, teacher implementation was unlikely before December, when LAPs were submitted. Because teachers would undertake their first assessments under the new system late in Term 1, the first questionnaire was circulated early in Term 2. All Round 1 questionnaires were completed before the end of June and corresponding interviews conducted over the next few months. Later rounds occurred approximately three months after the relevant quality management process. Guskey (2000) advocated a series of data collections “because implementation is often a gradual and uneven process” (p. 85). This recommendation was partly adopted. Because of concerns of excessive demands on participants, questionnaire items on assessment beliefs and practices were omitted from the second questionnaire.

### **3.3 Teacher invitation**

2011 provided a unique opportunity to investigate the potential for teacher learning in SACE Board quality management processes. For many subjects this entailed the first use of social moderation, with teachers having no experience of its use.

Investigating change in these teachers was an opportunity that would only exist in 2011.

It was decided to focus on teachers of one subject, Physics, a subject with no history of social moderation. Some subjects without previous social moderation might be taught by teachers with such experience in other subjects: Specialist Mathematics teachers might have taught Mathematical Applications. Language subjects had not used social moderation but had fewer teachers and, therefore, fewer potential participants. Other subjects, for example Tourism and Business Studies, had used social moderation previously (SACE Board of South Australia, 2010j), or had used group moderation, for example Biology and Chemistry. Physics satisfied the requirement of a large cohort of teachers with limited, or no, prior experience of social moderation.

### **3.4 Research methods and sources**

This section describes the various methods and sources used, the reasons for their selection, and the research questions that each addressed.

SACE Board declared itself “responsible for providing professional development programs for teachers” (SACE Board of South Australia, 2010h) and specified building “teacher understanding in the use of performance standards and in the design of assessment tasks” (SACE Board of South Australia, 2011j). The current research sought to identify elements in the quality management processes with the potential for teacher learning, particularly in teachers’ assessment beliefs and practices.

### 3.4.1 SACE Board documents

The first form of data identified were documents relating to the SACE Board quality management processes. Documents reveal “espoused theory” whereas interviews and observations reveal “theory-in-use”, a distinction observed by Patton (2002, p. 164).

Some SACE Board documents relating to the revised processes were available well before 2011. The earliest document used was the SACE Review (Crafter et al., 2006). While this was not produced by the SACE Board, it heralded the changes introduced. In December 2009 the moderation processes for 2011 were released (SACE Board of South Australia, 2010i). The February 2010 policy circular *Quality Assurance in the SACE* (SACE Board of South Australia, 2010f) and *Stage 2 Assessment and Quality Assurance in 2011* (SACE Board of South Australia, 2010k) of November 2010 were used in constructing research questions. The remainder of this section describes how the documents were used in subsequently addressing some of those questions.

These documents were read for evidence that the SACE Board quality management processes might display features described in the literature on effective professional development activities or effective teacher learning. Although evidence was also sought for features of effective adult learning and features that might encourage deep learning, it was anticipated that such evidence was more likely to be obtained by observation or participant interviews.

Second, documents were examined to identify content within the quality management processes that might contribute to teacher learning. The documents used in planning forums might, for example, include the reasons for the increased use of SBA and so influence teachers’ assessment beliefs. In such cases, it might be anticipated that teachers’ assessment practices might be changed as a consequence.

### 3.4.2 Observation of quality management processes

The second form of data collection used was observation of the quality management processes. Observations allow a researcher an understanding “not entirely possible using only the insights of others” (Patton, 2002, pp. 22, 23). Although observation has been described as “the most unbiased form of data collection [avoiding] ... the



self-report bias of surveys and interviews and allowing a clear look at what is actually occurring” (Desimone, 2009, p. 188), they have potential weaknesses: “a researcher may give meaning to ... [an] observation without checking out that meaning with participants” (Corbin & Strauss, 2008, p. 30). In the current research, subsequent interviews allowed observations and inferences to be refined.

Observations provide a view different from those of teachers, SACE Board officers and SACE Board documents. Thus the researcher has “less need to rely on prior conceptualizations of the setting” (Patton, 2002, p. 262) and is able to observe what might not be noted by a participant. The researcher can draw on personal knowledge and experience, therefore providing another lens from those of the participants.

Observations were undertaken of:

- a planning forum
- approval of schools’ LAPs
- training for conducting clarifying forums for teachers
- clarifying forums
- moderation

A decision was necessary concerning the nature of the observer’s role, which can vary from complete immersion (a participant) to complete separation (a spectator) (Patton, 2002). My role varied in different situations. Participant observation provides the opportunity “to see experiences from the views of participants” (Creswell, 2008, p. 222). Where this role was adopted, permission was sought and obtained from the person conducting the session.

In the planning forum, clarifying forums and moderation training, I attended as a participant—a subject teacher. In the LAP approval I had intended to be a spectator. The SACE Board officer, however, (aware of my SACE Board assessment experience) drew me into the group so that I undertook the training and participated in approving LAPs along with the other teachers.

A concern about observational data has been “the effects of the observer on what is observed” (Patton, 2002, p. 269). There is a concern that knowing they are under observation leads people to behave differently from if unobserved. Where I was a

participant this was not a concern. In LAP approval the panellists knew of my research but did not appear to be influenced by my presence; discussion appeared free-flowing. In the training for conducting clarifying forums, I was a spectator. There were few participants who knew that I was undertaking the research. They knew, however, of my similar role in clarifying forums in another subject. Because a participant observer is less conspicuous (J. P. Gall et al., 2005) and less threatening (Bogdan & Biklen, 1998) than an onlooker, it is argued that the influence on the conduct of anyone observed would be minimised, though this is not objectively verifiable.

Observations began with making notes on the settings and participants. During the session, notes were made on leader actions, participant actions, materials used and content covered. Reflective comments were added during and after the sessions. Observations were used to identify features that might contribute to the effectiveness of quality management processes as specified in the research sub-questions.

### **3.4.3 Questionnaires**

Questionnaires were used with the intention of collecting data from many teachers. Questionnaires are “lauded for being the only feasible mechanism for collecting data on large samples but are criticized for eliciting biased, socially desirable responses” (Desimone, 2009, p. 188) and were employed in the evaluation of moderation in Victoria for this purpose (Ingvarson, 1990). The price paid for asking many participants for many responses is that of standardisation— all are asked the same questions, which makes for easier tabulation and analysis.

Teacher recruitment used the SACE Board online forum for Physics teachers.<sup>16</sup> Online recruitment and questionnaire data collection allowed participation by a geographically diverse group of teachers. Although recruitment by this method might be criticised for favouring teachers with better access to online resources, it appears very few Physics teachers were not members of the online forum, as is now argued. The SACE Board annual report (SACE Board of South Australia, 2011b)

---

<sup>16</sup> Appendices include the invitations (Appendices A and B), letters of introduction (Appendices C and D), the research information sheet (Appendix E) and the consent forms (Appendices F and G).

stated that 117 schools offered Stage 2 Physics and 147 schools submitted moderation samples (SACE Board of South Australia, 2011a). This difference arises from classes, in different schools, moderated separately combining for the written examination. There were 181 classes moderated (SACE Board of South Australia, 2011a). It would appear that the number of South Australian Stage 2 Physics teachers were in the range 147–181. The online forum had 271 members in 2011 (SACE Board of South Australia, 2011c), suggesting most Physics teachers were online forum members. It was hoped that a large number of Physics teachers would be willing to participate in the questionnaires. Data from many participants might allow generalisation about the population under study—Physics teachers of South Australia. Nevertheless, it is acknowledged that such data would be broad-brush since an online questionnaire cannot identify and describe nuances in participants' positions. The questionnaires included an invitation to be interviewed.

Questionnaires included multiple-choice and open-ended items: “an example of how quantitative measurement and qualitative enquiry are often combined” (Patton, 2002, p. 5). Multiple-choice questions employ “a deductive approach because items must be predetermined” (Patton, 1987, p. 15); many of these were based on a review of the literature. Open-ended responses permitted participants to “describe what [was] meaningful and salient without being pigeon holed into standardized categories” (Patton, 2002, p. 56). These items provided “more flexibility and greater latitude” (Guskey, 2000, p. 101), yielding idiosyncratic information about teachers' opinions, which was explored in subsequent interviews. The open-ended items took the form: “Please add any other comments you wish to make regarding ...” and responses treated as part of the corresponding interview.

Questionnaire construction was based on a number of instruments used in researching teacher learning and attitudes and practices related to assessment. Attempts were made to avoid double-barrelled items, use short and simple items, avoid negatively stated items and avoid biased items or terms (Babbie cited in DeLuca & Klinger, 2010, p. 426). Similar recommendations were made by Creswell (2008): “cut out unnecessary words”, avoid jargon, avoid “multiple questions” and avoid questions that are not “applicable to all participants” (p. 402).

To address the many facets, the questionnaires were long, raising concern that their

length might dissuade participants. Consequently the questions of most interest to teachers were placed early and the easily answered questions about personal information were placed at the end (Fink, 2009). It is acknowledged that positioning these questions near the end was counter to advice that “questions should proceed from the most to the least familiar” (Fink, 2009, p. 37).

For the multiple choice items a 5-point Likert scale was adopted: strongly disagree, disagree, neutral, agree, strongly agree. Although a 4-point scale forces respondents to agree or disagree, neutral being unavailable (Fink, 2009), a 5-point scale was used because, for many items, teachers might resist a disagree–agree decision and “forcing respondents to choose may annoy them and may not uncover the truth about their views” (Fink, 2009, p. 26). Numerous agree-disagree items made it impossible to comply with advice to “avoid many items that look alike” (Fink, 2009, p. 37). Nevertheless her recommendation to group questions was adopted in the hope of minimising loss of interest.

A questionnaire pilot was conducted with three teachers who were invited to comment on the questionnaire, leading to modifications. Pilot volunteers were not SACE Physics teachers, thus avoiding the need to exclude members of the pilot group from those participating in the study (Creswell, 2008). The modified questionnaire was then used online with six Chemistry teachers and feedback used to amend the questionnaire and the online procedure. In the pilot, the centre point on the scale had been labelled “No opinion”; some respondents indicated that they had an opinion, but that it was neutral. Consequently, this was amended to “Neutral”. Changes were made to the wording of some items where the teachers reported clarification was necessary.

The questionnaires were used to construct participant profiles in terms of the teacher characteristics identified earlier: professional, school and personal. Data were collected in six categories: assessment beliefs, assessment practices, school support, experience of SACE Board activities, personal information and learning about assessment. Each dimension will now be described.<sup>17</sup> The section on questionnaires concludes with an overview of the data collected in each round and the research

---

<sup>17</sup> The questionnaires are provided in Appendices H–J.

questions addressed by each dimension.

### **3.4.3.1 Assessment beliefs**

The connection between beliefs and practices is disputed (e.g., James & Pedder, 2006a; Richardson & Placier, 2001), with disagreement about whether changes in beliefs and values precede (e.g., Fullan, 2007) or follow changes in practices (e.g., Guskey, 2000). Although the SACE Board quality management processes aimed to enhance teachers' assessment practices (SACE Board of South Australia, 2010f), the current research also investigated teachers' beliefs—first to see if they might be associated with teacher involvement in the quality management processes and, second, to seek evidence of change in these beliefs.

The questionnaires addressed four aspects of teachers' assessment beliefs. The first, *confidence in formative*, examined their confidence in SBA “for creating student learning opportunities in the classroom” (James & Pedder, 2006a, p. 112). The second, *confidence in summative*, considered teacher confidence in the summative use of SBA, particularly as practised by the SACE Board. The third aspect, *SACE moderation knowledge*, addressed teachers' beliefs about the purposes of social moderation while the fourth, *SACE self-confidence*, focussed on teacher confidence in the assessment practices required of them by the SACE Board. Discussions of each of these aspects will be followed by an overview of the assessment beliefs section of the questionnaires.

#### **3.4.3.1.1 Confidence in formative**

The terms *formative* and *summative* are commonly used to describe different assessment practices. Gardner (2006) traced the term “formative assessment” to Scriven, who used the terms to discriminate “solely on the basis of when the evaluation” (p. 5) was carried out. More recently the terms have been used to describe “the use to which the assessment is put” (William, 1992, p. 18). Thus assessment is described as formative when “used to adapt the teaching work to meet the students' learning needs” (Black, Harrison, Lee, Marshall, & William, 2003, p. 2) while summative assessments “count towards a final result” (Mercurio, 2008, p. 3).

James and Pedder (2006a) investigated teachers' values regarding formative assessment practices and the degree to which their practices complied with these

values. In the current research it was considered that the value placed on formative assessment might be associated with willingness to accept the assessment changes of the new SACE. Consequently data on this dimension were collected in the questionnaires.

Because the questionnaire aimed to address more areas than James and Pedder (2006), the decision was made to limit the number of items used. Such statements were not relevant to the current research concerning written summative SACE assessment; for example “I use questioning mainly to elicit reasons and explanations from my students” (Item A18, p. 136). Others were reworded. For example, the item “Assessment of students’ work consists primarily of marks and grades” (Item A12, James & Pedder, 2006a, p. 137) became “Assessment of student work should always be reported as marks.” This was one of the statements where some change in responses was anticipated.

The pilot identified items similarly scored by all participants, for example “The new assessment requirements are too time-consuming”. Teachers in the pilot commented that some items would receive similar responses from all participants and that responses were unlikely to change over a 12-month period. Consequently, such items were omitted from the eventual questionnaires.

Although Likert scales were used in the current research, the format differed from that of James and Pedder (2006a). They had used the same items for investigating teachers’ values and practices, using two adjacent grids as shown in Figure 8.

Scale X				Section A	Scale Y				
This school now					Teachers’ assessment practices	How important are your assessment practices for creating opportunities for students to learn?			
(About you)				(About your values)					
Never true	Rarely true	Often true	Mostly true		Not at all important	Of limited importance	Important	Crucial	Bad practice
				I provide guidance to help my students assess their own work.					

**Figure 8: Dual scale format used in James and Pedder’s teacher questionnaire**

Because the current research did not attempt to address coherence between teachers’ beliefs and practices, identical items for beliefs and practices were unnecessary. Using two grids in the current research allowed rewording of some statements to

more clearly articulate beliefs and practices. Thus, for example, James and Pedder's (2006a) statement "I provide guidance to help students assess their own work" (Item A13, p. 120) was amended to "Students should be given guidance to assess their own work" in the beliefs section, while the initial wording was retained for the practices section. Table 7 lists James and Pedder's (2006a) statements used in the current research to gather data on teachers' beliefs about the formative use of SBA.<sup>18</sup>

**Table 7: Items related to teacher confidence in the formative use of SBA**

The prescribed curriculum is more important for planning lessons than how well students did in the last lesson
The most useful assessment for student learning is assessment that is undertaken by the teacher rather than an external examination
Assessment should help students to learn independently
Students should be given guidance to assess their own work
Students should be given guidance to assess other students' work
Students benefit from opportunities to assess other students' work
Assessment of student work should always be reported as marks
Student effort is irrelevant when assessing their work
It is important that students are told how they performed in relation to others in the class

Items from Table 7 were used to generate scores for a derived variable, *confidence in formative*. The decision was made to treat the Likert-scale responses as continuous data, in which case "means and standard deviations are appropriate statistics" (Fink, 2009, p. 26). The sum of responses was analogous to the process described by Gall et al. (2005). Calculation of the means (to one decimal place) was used to generate a 0–5 scale for teachers' *confidence in formative*. Similar scales were generated for other aspects of teachers' assessment beliefs and practices. De Vaus (2002) advocated the development of a "composite measure of a concept" (p. 180) because "it helps get at the complexity of the concept" (p. 181). He described how multiple indicators provide more valid measurement, increase reliability and precision, and simplify analysis. Although "a single scale ought to measure a single construct" (Briggs & Cheek, 1986, p. 109), there is the problem of "how to specify the notion of a single construct both conceptually and operationally" (Briggs & Cheek, 1986, p.

<sup>18</sup> Appendix R shows the relationships between the questionnaire items and those of James and Pedder (2006a).

110). The first stage of this process was selecting items that appeared to measure the same construct: “Most of the effort to ensure validity takes place before any data gathering” (Nunnally & Bernstein, 1994, p. 304). Items that James and Pedder (2006a) had considered as *promoting learning autonomy* (p. 122), for example, were treated similarly. Items identifying the purpose of assessment as measurement of learning outcomes and ranking of students, rather than as part of the learning process, were reverse-coded. The first item in Table 7, for example, emphasises the curriculum as determining what is taught rather than student learning and so was reverse-coded, as was the item on the use of marks, which is contrary to providing feedback that “focuses students on specific strategies for improvement” (Brookhart, Moss, & Long, 2010, p. 41).

The expectation of “at least a moderate level of internal consistency among the test items; i.e., the items should tend to measure something in common” (Nunnally & Bernstein, 1994, p. 103) can be judged with participant data. Reliability was assessed using Cronbach’s alpha, “the most widely used objective measure of reliability” (Tavakol & Dennick, 2011, p. 53), which is “a function of internal consistency, that is, of interrelatedness of items” (Cortina, 1993, p. 100). Acceptable values for alpha vary “from 0.70 to 0.95” (Tavakol & Dennick, 2011, p. 54). The data were approached with caution because of the small sample size and the small number of items, both of which reduce reliability (Cortina, 1993; Nunnally & Bernstein, 1994; Tavakol & Dennick, 2011). Reliability also depends on the mean inter-item correlation (Nunnally & Bernstein, 1994), recommended as “an alternative index of internal consistency” in Cronbach’s initial paper on alpha (Green, Lissitz, & Mulaik, 1977, p. 834) and by others (De Vaus, 2002; Nunnally & Bernstein, 1994; Russell & Russell, 1994; Tavakol & Dennick, 2011). Briggs and Cheek (1986) suggested mean inter-item correlations should be in the range 0.2 to 0.4 (p. 115). Lower inter-item correlations suggest “each part of the test must be measuring something different” (Kline reported in Boyle, 1991, p. 291) while higher correlations suggest there is some redundancy (Briggs and Cheek reported by Piedmont & Hyland, 1993).

The questionnaire data were analysed using SPSS. Cronbach’s alpha for the nine items described above was 0.172 with a mean inter-item correlation of 0.030.



Deletion of the first two items and the item on always reporting student work as marks generated a Cronbach's alpha of 0.502 and a mean inter-item correlation of 0.169, neither of which was in the preferred range<sup>19</sup>. Nevertheless, it was felt that omitting more items would leave too few for a valid representation of teacher confidence in the formative use of SBA. *Confidence in formative* data were used to analyse whether differences were associated with involvement in the quality management processes and whether change occurred over the research period.

#### **3.4.3.1.2 Confidence in summative**

Data were also collected on teachers' beliefs about the summative use of SBA, particularly SACE Board assessment. These data were used to generate scores for a second derived variable, *confidence in summative*. Twelve statements were written about the reliability and validity of SBA for external assessment, with items worded using SACE Board terms.

It seemed that teachers would want SBA to be used for summative purposes if they believed it could assess learning outcomes, such as practical and collaborative skills, that could not be assessed by examinations. Teachers who saw performance standards as providing a framework for formative assessment might also see benefit in using them for summative purposes. Conversely, teachers believing, for example, that the LAP served no real purpose, were unlikely to have confidence in using SBA for summative purposes. Thus, four of the items were:

- SBA can assess more learning outcomes than external assessment.
- A learning and assessment plan is merely bureaucratic; it serves no real educational purpose.
- The performance standards are the foundation for designing assessment tasks.
- Student work can be graded without the SACE performance standards.

The questionnaires included items that provided opportunities for teachers to express doubts about the suitability of SBA for summative purposes because it was less rigorous, subjective or open to manipulation:

---

<sup>19</sup> A more detailed discussion of the questionnaire data analysis is presented in Appendix R.

- SBA is less rigorous than external assessment.
- Students have too many opportunities to receive help with SBA.
- SBA is open to manipulation by teachers.
- I am not confident that other teachers make sure that the work they mark is the students' own work.
- External assessment is the fairest form of assessment.
- External assessment is objective.
- I believe that some teachers will mark generously because moderation will not make small changes to teachers' grades.

There was also an item suggesting that the summative use of SBA placed teachers in an invidious position:

- There is a role conflict when a teacher is required to be an assessor.

Thus there were twelve items. Inadvertently the items had been expressed in forms with which teachers might agree; it was later realised that this was contrary to the advice to avoid items that look alike (Fink, 2009) and that items be worded to generate varying responses (Rosenholtz, 1989). The large number of negatively worded items may have influenced responses. Items considered to express an unfavourable opinion of the summative use of SBA were reverse-coded.

Cronbach's alpha for the twelve items was 0.385 with a mean inter-item correlation of 0.043. Analysis using SPSS identified nine items<sup>20</sup> giving a Cronbach's alpha of 0.696 and a mean inter-item correlation of 0.229. Means were calculated to give scores out of five for *confidence in summative* that were used to analyse for association with involvement in quality management processes and to see whether change occurred over the research period.

#### **3.4.3.1.3 SACE moderation knowledge**

The third aspect of teachers' beliefs concerned social moderation. It was considered that teacher knowledge might change during the research period and that teachers with greater involvement in the quality management processes might show greater change. Table 8 shows the items used for this purpose.

---

<sup>20</sup> Appendix R shows the processing of questionnaire responses in greater detail.

**Table 8: Items related to teachers' beliefs about the purposes of moderation**

Moderation is necessary to ensure that all teachers have followed the subject outline in their assessment
Moderation is necessary to ensure that all teachers have followed their learning and assessment plan
Moderation is necessary to ensure that students have undertaken comparable assessment tasks under similar conditions
Moderation is necessary to ensure that students have not received undue help in assessment tasks
Moderation is necessary to ensure that teachers have marked according to the performance standards

No aggregation was used for these items. Adding or averaging would suggest that a higher score, indicating agreement, shows greater awareness of the purposes of social moderation and that correlation between responses might be expected. That this is incorrect is demonstrated by an example. In Round 1 a participant might strongly agree that moderation is undertaken to ensure that teachers have followed their LAPs. This person might observe that the LAP was not used in moderation, and subsequently disagree with the statement. Thus, although the teacher is more informed about social moderation, the score would have decreased. The analysis was, therefore, limited to comparing Rounds 1 and 3 responses to individual items to determine whether a teacher's knowledge might have changed over the research period.

#### **3.4.3.1.4 SACE self-confidence**

The fourth aspect of teacher assessment concerned confidence of teachers in their own assessment practices as required by the SACE Board: *SACE self-confidence*. It was anticipated that teacher confidence would develop during the research period and might be greater for teachers participating more in the quality management processes. Table 9 shows the items used to collect these data.

**Table 9: Items related to confidence of teachers in their own assessment practices**

I believe that my next learning and assessment plan will be approved without amendment
I believe that my assessment tasks address all of the specific features of the subject outline
I believe that my grading of student work is aligned with the standards of other teachers
I believe that my interpretation of the performance standards is aligned with that of other teachers
My assessment tasks allow students to demonstrate work at an A+ grade level

Averaging the five responses generated *SACE self-confidence* scores. Correlation data were not determined for two reasons: the items were believed to be independent of each other and preclusion of some would diminish the data gathered about confidence of teachers in their own assessments.

*SACE self-confidence* data were used to analyse whether they were associated with involvement in quality management processes and whether change occurred over the research period.

#### **3.4.3.1.5 Assessment beliefs—overview**

All four aspects of teachers' assessment beliefs were included in Round 1 and Round 3 questionnaires. Data collected on teachers' assessment beliefs in Round 1 might be associated with involvement in the quality management processes and could be used in interviewee selection. It will be explained later why this was not done (see 3.8). Comparisons of data from Rounds 1 and 3 might suggest changes in teacher assessment beliefs during the research period.

#### **3.4.3.2 Assessment practices**

The second dimension of the questionnaires concerned teachers' assessment practices since they were the focus of the quality management processes undertaken by the SACE Board (SACE Board of South Australia, 2010f). Although the SACE Board addressed summative practices, it was considered that some changes in formative practices might occur. Consequently, data were collected on formative and summative practices. These two aspects will now be described, followed by an overview of the sections about teachers' assessment practices.

### 3.4.3.2.1 Formative practice

James and Pedder's (2006a) research into classroom assessment was referred to earlier (see 3.4.3.1) where it was indicated that they used the same items for beliefs and practices. This approach was not adopted in the current research, which used eleven of their statements for collecting data relating to teachers' assessment practices. Some statements were used in their original form, whereas other items were reworded to make them more relevant to SACE Physics teachers. Table 10 shows the items which collected data on teachers' assessment practices creating learning opportunities for students.<sup>21</sup>

**Table 10: Statements related to teachers' practices in the formative use of SBA**

I adjust my teaching plan depending on how well students understood the work completed in class
My assessment of students' work consists primarily of marks
My assessment practices help students to learn independently
I provide guidance to help students assess their own work
I provide guidance to help students assess other students' work
Students are given opportunities to assess other students' work
I encourage students to view mistakes as valuable learning opportunities
I use students' misconceptions shown in their work to adjust my teaching
My assessment of student work includes extended written comments
I use my knowledge of a student's effort when I assess their learning to determine their final mark
When returning student work I tell students how well they have done in relation to others in the class

Treating the responses as continuous data, means were used to generate scores for *formative practice*. Reverse-coding was used for the second and last items. Cronbach's alpha was 0.278 and the mean inter-item correlation was 0.069. Omitting the reverse-coded items gave an alpha value of 0.718 and a mean inter-item correlation of 0.241 and the remaining items were used to generate *formative practice* scores.

*Formative practice* data were used to analyse whether differences were associated

<sup>21</sup> Appendix R shows the relationships between the questionnaire items and those of James and Pedder (2006a).

with involvement in quality management processes and whether change occurred over the research period.

#### 3.4.3.2.2 Summative practice

Additional items were written that sought information on teachers' summative assessment practices. Table 11 shows these statements.

**Table 11: Statements related to teachers' summative assessment practices**

I am able to design effective assessment tasks without referring to the assessment design criteria
I find it hard to assess my students' work because my main role as a teacher is to help them develop educationally, emotionally and socially
I make sure that my marking of each task is against the performance standards
I make sure that I apply the performance standards in the same way to all students in my assessment
When assessing I tend to be generous because I know that moderators will not be allowed to make small changes to students' grades
When assessing I tend to be a bit harsh because I do not want to risk my grades being moderated downwards

Analysis gave a value of 0.018 for Cronbach's alpha and a mean inter-item correlation of -0.002. Omitting the first item and reverse-coding the last generated a value of 0.382 for alpha and a mean inter-item correlation of 0.123.<sup>22</sup> Thus scores were generated for *summative practice*.

*Summative practice* data were analysed for association with involvement in quality management processes and whether change occurred over the research period.

#### 3.4.3.2.3 Assessment practice—overview

Round 1 assessment practices data contributed to the personal characteristics intended for interviewee selection. The data were scrutinised for any association between teachers' assessment practices and involvement in the quality management processes. Comparison of data from Rounds 1 and 3 might suggest changes in teachers' assessment practices during the research period.

<sup>22</sup> Appendix R shows the processing of questionnaire responses in greater detail.

### **3.4.3.3 School support**

The third dimension of the questionnaires, on school support, collected data on four aspects of the school environment: goal-setting, shared goals, collaboration, and learning opportunities. It also collected data on the schools' expectations regarding involvement in SACE Board activities and support provided for such involvement.

The current research was based on the premise that teachers who chose to be involved in the quality management processes expected to learn from the involvement. Because the effectiveness of teacher learning is strongly linked to school culture (Guskey, 2000; Loucks-Horsley et al., 2003; Rosenholtz, 1989; Timperley, 2008), participants were asked about variables that Rosenholtz (1989, p. 9) used to identify schools as learning-enriched or learning-impooverished. It was anticipated that the culture of a teacher's school might be associated with involvement in the quality management processes. Since "school culture" might be viewed negatively by participants, the term "school expectations" was adopted in the questionnaires.

The exploratory nature of the current research meant that the data on school support constituted only one element of the data to be collected. A belief that it was desirable to limit the questionnaire length restricted the number of items addressing school support. Consequently, from the 164 items Rosenholtz used, 20 were selected, being those considered most appropriate to stand-alone usage, relevant to the SACE context and giving variation in negative and positive wording. Table 12 shows the items used, grouped according to her organisers: *goal-setting*, *shared goals*, *collaboration* and *learning opportunities*.

**Table 12: Items related to school expectations and support, from Rosenholtz (1989)**

Goal-setting
Discussion about school goals and means of achieving them is a regular part of our staff and faculty meetings
There are a lot of irrelevant discussions at our faculty meetings
At faculty meetings, we spend most of our time on the small stuff; we rarely get a chance to talk about the bigger issues in teaching and learning
The principal of my school encourages teachers to talk with each other about instructional objectives
Shared goals
I don't approve of the ways in which most of the other teachers in this school teach
Most teachers at my school have values and philosophies of education similar to my own
Teachers at this school share a high level of commitment to student learning
My principal's values and philosophy of education are similar to my own
Collaboration
I don't offer advice to others about their teaching unless I am asked for it
Other teachers at this school come to me for help or advice when they need it
I can get good help or advice from other teachers at my school when I have a teaching problem
I give help and support to other teachers when they are having problems in their teaching
Other teachers at this school seek my advice about professional issues and problems
I regularly share teaching ideas or materials with other teachers
Learning opportunities
New ideas presented at in-services are discussed afterwards by teachers in this school
At this school, I have many opportunities to learn new things relevant to my teaching
I receive informal evaluations of my teaching performance from other teachers
Other teachers encourage me to try out new ideas
My principal encourages me to try out new ideas
When teachers are not doing a good job, our principal works with them to improve instruction

Items where agreement was interpreted as a negative view of the school's culture, such as irrelevant discussions, were reverse-coded.<sup>23</sup> Means were calculated for each group, generating values for *goal-setting*, *shared goals*, *collaboration* and *learning*

<sup>23</sup> The reverse-coded items are identified in Appendix R.



*opportunities*.<sup>24</sup>

Because the current research was interested in teacher involvement in SACE Board quality management processes, other items were written for *school support for SACE*. Table 13 shows the items used for collecting these data.

**Table 13: Items related to school support for SACE Board involvement**

My school encourages attendance at SACE Board teacher meetings
My school encourages teachers to act as SACE Board markers or moderators
Attendance at SACE Board teacher meetings is not encouraged in my school unless relief teachers are provided by the SACE Board
Teachers in my school are not able to act as SACE Board moderators unless relief teachers are provided by the SACE Board

The last two were reverse-coded since agreement suggested lack of encouragement for participation in SACE Board processes. Reliability analysis gave a value for Cronbach's alpha of 0.767 and a mean inter-item correlation of 0.489. The mean was used as the value for *school support for SACE* and the mean of this and the four values of Table 12 gave values for *school support*. These data were only collected in the Round 1 questionnaire and were part of the teachers' school characteristics for interviewee selection and exploring possible associations with involvement in the SACE Board's quality management processes.

#### **3.4.3.4 Experience of SACE Board quality management processes**

The fourth dimension of the questionnaires sought data on teacher involvement in, and experiences of, the SACE Board quality management practices. The teacher-SACE Board relationship was enacted through published documents, the Board website and the quality management processes. The current research was implemented on the premise that it was through these means that the SACE Board addressed its responsibility for teacher professional development (e.g., SACE Board of South Australia, 2010h; SACE Board of South Australia, 2011j). Data on teachers' experiences of the quality management processes were an important part of

---

<sup>24</sup> Because these items were from previous research (Rosenholtz, 1989), no consideration was given to omitting any. Nevertheless, data were analysed for reliability with results shown in Appendix R.

the current research and were collected in each round of the questionnaires.

In Round 1, data were collected on teachers' experiences in planning forums and LAP approval, and about LAP feedback. In Round 2, data were collected on clarifying forums, while Round 3 data concerned moderation experiences and moderation feedback. Items addressed four possible areas of learning: writing a LAP, designing assessment tasks, using the performance standards in marking student work, and applying the performance standards in a manner consistent with other teachers. Responses were on the 5-point Likert scale described previously.

These data were used for evidence that the quality management processes might have contributed to teacher learning and to frame interview questions.

#### **3.4.3.5 Teacher information**

Teacher information was collected in the Round 1 questionnaire and used in generating teacher profiles. Gender data contributed to participants' personal characteristics. The highest level of study in Physics, years of teaching, years of teaching Stage 2 Physics and previous participation in SACE Board assessment panels contributed to the professional characteristics. School location (metropolitan or regional South Australia) and the educational sector of their school (Catholic, Government, or Independent) contributed to their school characteristics.

These data were for interviewee selection and to explore whether teacher involvement in the quality management processes might be associated with teacher characteristics.

#### **3.4.3.6 Learning about assessment**

The sixth dimension of the questionnaires, the final part of the Round 3 questionnaire, investigated the depth of possible learning during the research period. Hattie (2003) made the distinction between surface learning, more about content, and deep learning, more about understanding. Mezirow's transformation theory relates learning to "making meaning" (Mezirow, 1994, p. 222) and identifies eleven phases in the learning process. *The Learning Activities Survey Questionnaire* (Brock, 2010), used to identify learning steps among undergraduate students, was modified so that the items more clearly related to learning that might be associated with quality

management processes. Table 14 shows the items as used by Brock, the modified forms used in the current research, and two additional items that emerged from interviews as teachers described their learning during the year.

**Table 14: Items used to identify transformative learning phases**

<b>Brock's items</b>	<b>Current research items</b>
I had an experience that caused me to question the way I normally act.	I have questioned my beliefs about assessment.
I had an experience that caused me to question my ideas about social roles.	I have questioned the way I carry out assessment.
As I questioned my ideas, I realized I no longer agreed with my previous beliefs or role expectations.	I realized I no longer agreed with my previous beliefs.
Or instead, as I questioned my ideas, I realized I still agreed with my beliefs or role expectations.	I realized I still agreed with my original beliefs.
I realized that other people also questioned their beliefs.	I realized that other teachers were also questioning their beliefs about assessment.
I thought about acting in a different way from my usual beliefs and roles.	I thought about carrying out my assessment in a different way.
I felt uncomfortable with traditional social expectations.	I felt uncomfortable with the previous form of assessment.
I tried out new roles so that I would become more comfortable or confident in them.	I tried out new forms of assessment so that I would become more comfortable or confident with them.
	I tried out new forms of assessment so that I could select the best format.
I tried to figure out a way to adopt these new ways of acting.	I tried to figure out a way to adopt new ways of assessing.
I gathered the information I needed to adopt these new ways of acting.	I gathered information I needed to adopt new ways of assessing.
I began to think about reactions and feedback from my new behaviour.	I began to think about reactions and feedback about my new forms of assessment.
I took action and adopted these new ways of acting.	I took action and adopted new ways of assessing.
	I changed my ways of assessing a number of times.

These data were used in interview preparation and as evidence that involvement with the SACE Board's quality management processes might have contributed to teachers

learning about assessment.

### 3.4.3.7 Overview of teacher questionnaires

As mentioned earlier, data were collected at three points in the research period, referred to as three *rounds* of data collection. Table 15 shows the sequence of data collection.

**Table 15: Data collected in each round of questionnaires**

	Round 1	Round 2	Round 3
Assessment beliefs	✓		✓
Assessment practices	✓		✓
School support	✓		
<b>Experiences of quality management processes</b>			
• Planning forum	✓		
• Approving learning and assessment plans	✓		
• Feedback on learning and assessment plans	✓		
• Leading a clarifying forum		✓	
• Attending a clarifying forum		✓	
• Serving as a moderator			✓
• Feedback after moderation			✓
Personal information	✓		
Invitation to be interviewed	✓		
Learning about assessment			✓

Table 16 shows the research questions addressed by the data collected in the questionnaires. Round 1 occurred early in Term 2 2011, Round 2 early in Term 4 2011 and Round 3 in the middle of Term 1 2012.

**Table 16: Questionnaire dimensions and research questions addressed**

<b>Dimension</b>	<b>Research Question 1</b>	<b>Research Question 2</b>	<b>Research Question 3</b>	<b>Research Question 4</b>
Assessment beliefs	✓			✓
Assessment practices	✓			✓
School support	✓			
Experiences of quality management processes	✓	✓	✓	✓
Personal information	✓			
Learning about assessment				✓

### 3.4.4 Teacher interviews

This section describes the reason for including interviews in the research design. It then explains the interviewee selection process and the development of the interview protocols.<sup>25</sup>

To address the deficiency of standardisation implicit in questionnaires (Patton, 2002), interviews were conducted to “elicit data of much greater depth than is possible with other measurement techniques” (J. P. Gall et al., 2005, p. 134). It was hoped that teachers would respond to the invitation in the first questionnaire to participate in a series of one-hour interviews over the year, providing the opportunity to “explore changes in specific individuals and the possible reasons for such changes” (J. P. Gall et al., 2005, p. 182).

A strategy was required for selection of interviewees from those willing to be interviewed. With a small sample, it was unlikely that the interviewees would be representative of the population. The intent to undertake multiple interviews with each participant, however, suggested a small sample as appropriate in order to make the task manageable. It was suspected that professional characteristics (e.g., teaching experience) might influence involvement in, and response to, the new quality management processes. School characteristics (e.g., school support) and personal characteristics (e.g., assessment beliefs) might also affect involvement.

Consequently, it was decided to employ stratified purposeful sampling which, by selecting identified subgroups, “facilitates comparisons” (Punch, 2006, p. 51). Thus

---

<sup>25</sup> The interview protocols are provided in Appendices K–M.

the criteria for selecting interviewees related to teachers' professional, school and personal characteristics. Interviewees were to be selected to provide a purposive sample which, while not representative, contained members possessing a range of characteristics. Given the gender imbalance among Physics teachers, gender balance was considered unlikely. Earlier research (B. Atherton, 2009) showed that, with a similar purposive sample, new information did not appear after interviewing fifteen teachers. Consequently, to allow for some loss of participants during the research period, it was decided to begin with twenty interviewees from diverse teaching situations, with differences in teaching history and SACE Board assessment experience.

An interview may be one-to-one or a group process. In group discussions participants may become more aware of their attitudes than when interviewed alone, or their attitudes may change during the discussion. There is also a possibility that, because of a perceived need to "belong", participants might subscribe to others' opinions that they do not hold. While a group interview would facilitate gathering the opinions of more people, this form was rejected because of the greater likelihood of teachers giving their own opinions if interviewed individually.

Interviews may be open, partly structured, highly structured or use a pre-coded format (Foddy, 1993). Open interviews were rejected because of the need to focus discussion on the quality management processes. In a structured interview, the interviewer asks each participant exactly the same questions with responses limited to a set provided to participants and recorded using a pre-established coding scheme (Fontana & Frey, 2000). This format was rejected as too inflexible because it did not allow for unexpected responses.

Since the research was exploratory, it was necessary to address issues that emerged in the questionnaires and during the interviews. It was therefore decided to use a sequence of semi-structured interviews, as used in earlier research into teacher learning (B. Atherton, 2009; Fastier, 2007; Fernandez, 2007; L. Reid, 2007). Interview protocols<sup>26</sup> were developed listing questions to be addressed in each interview. These were open-ended, allowing participants to frame their responses,

---

<sup>26</sup> See Appendices K–M.

with clarification to be sought where necessary. Each topic was introduced with a lead question, with additional questions when clarification was necessary. Interview preparation entailed selection and modification of questions in the light of responses to the associated questionnaire, providing richer data than from a structured interview. References to questionnaire responses were considered beneficial in developing “a trusting relationship between the interviewer and interviewee” (Desimone, 2009, p. 188).

To design interviews that avoided ambiguous questions and responses, an initial protocol was used with three teachers who had piloted the questionnaires who were not participants in the current research. The ensuing discussions improved question clarity and removed ambiguities. Minor modifications were made during the research period as necessary.

Interviewers may affect participants’ responses: “interviewer bias” (Desimone, 2009, p. 188). Interviews are not neutral tools; the inter-personal interactions generate “negotiated, contextually based results” (Fontana & Frey, 2000, p. 646). What is said in interviews is shaped by the interviews themselves and they have been perceptively described as “collaborative accomplishments” (Holstein & Gubrium, 2004, p. 141).

Another disadvantage of interviews arises with multiple interviewers reducing reliability of the responses. Although interviewer characteristics have “relatively small impact on responses” (Fontana & Frey, 2000, p. 646), having one interviewer eliminated different interviewers with different effects on participants.

To minimise interviewer bias, questions were asked and minimal comments made on responses. While neutrality is important, interviewer unresponsiveness risks the climate of trust which encourages an interviewee to talk more easily (Rapley, 2004). A balance was needed. While it was necessary not to influence the interviewee's story and “contaminate” the data (Rapley, 2004, p. 19), I was cautious of being seen as aloof, with the interviewee feeling like “a research object” (Rapley, 2004, p. 19). In good interviews “the subjects are at ease and talk freely about their points of view” (Bogdan & Biklen, 1998, p. 95). Indeed, some have argued that researchers should belong to the groups being studied, “in order to have the subjective

knowledge necessary to truly understand their life experiences” (Miller & Glassner, 1997, p. 105). Because of my similar background I was able to establish rapport with the teachers interviewed. In some cases my science teaching experience provided the bond; in others, my experience as a marker while, for a few, SSABSA and SACE Board experiences provided the connection. Familiarity with their situations meant there were few occasions when clarification was needed. I was able to empathise with them, so that elaboration was forthcoming when sought.

All interviews were recorded with participant agreement. After transcription, each participant was provided with the transcript and advised of the right to correct the transcript and ask for parts to be excluded. While some participants clarified, none requested any exclusions. The verified transcripts were imported into NVivo for analysis.

The interviews were critical in the research design, providing data on many aspects of teachers’ experiences of the 2011 SACE Board quality management processes. Some processes (e.g., preparing materials for moderation) were addressed once while others (e.g., social moderation beliefs) were addressed more often. To identify changes in teachers’ beliefs or practices would require a number of interviews for each teacher. There was concern that teachers might not be willing to commit to a large number of interviews. Consequently, some aspects were grouped together so that a participant might be interviewed six times during the research period: three times in Round 1, once in Round 2 and twice in Round 3. The interviews addressed assessment experiences, verbalisations of assessment practices and critiques of their assessment tasks. The topics addressed in each interview are shown in Table 17. Each is described, and the reasons for their inclusion explained, in the following sections.



**Table 17: Overview of interview topics**

<b>Dimension</b>	<b>Round 1</b>	<b>Round 2</b>	<b>Round 3</b>
<b>Assessment experiences</b>	✓	✓	✓
Experiences of quality management processes	✓	✓	✓
Learning about assessment	✓	✓	✓
Assessment beliefs	✓		✓
Assessment practices	✓	✓	✓
<b>Verbalisation of assessment practices</b>	✓		✓
<b>Own task critique</b>	✓		

### **3.4.4.1 Assessment experiences**

This dimension addressed four topics: (1) teachers' experiences of the SACE Board quality management processes; (2) teachers' reflections on learning about assessment; (3) teachers' assessment beliefs and (4) teachers' assessment practices.

#### **3.4.4.1.1 Teachers' experiences of the quality management processes**

Participant experience of the quality management processes was addressed in each round of interviews. Each time participants were asked what they believed the SACE Board hoped to achieve in the process, how well those aims had been achieved and the outcomes for the participants. Scrutiny of individual questionnaire responses was used to modify questions prior to the interview. Thus, for example, one interviewee was asked to expand on the comment in which the planning forum was described as "too late and leaving a lot of questions open". The participant who commented that the clarifying forum "raised more questions than it answered" was reminded of the comment and asked for some of those questions.

Round 1 collected data on experiences of the planning forums and LAP approval. Two participants had been involved in LAP approval while others were asked about feedback from the LAP approval process. Round 2 interviews addressed the clarifying forums: all had attended as subject teachers, while two also served as presenters. In Round 3 all participants were asked about moderation feedback while six reflected on their experiences as moderators.

In each case, participants were asked to identify and describe the opportunities for teacher learning in the SACE Board quality management processes and reasons for their involvement.

#### **3.4.4.1.2 Teachers' reflections on their learning**

The second topic of participants' assessment experiences sought information on how they had learned about assessment, but did not directly refer to the quality management processes. Their purpose was to stimulate teachers to reflect on the sources they had used in making changes to their assessments during the year. Bissaker (2009) described teachers as failing to recognise skills and knowledge acquired in their "work" as learning. Similarly, teachers may not recognise the quality management processes as learning. When asked where they had learned about assessment, however, it was possible that they might give more credit to the quality management processes than if asked directly.

The Round 1 interview included teachers' reflections on previous assessment learning, as well as learning associated with the changes of 2011. The lead question was: "I would like to ask you to look back over your time as a teacher and describe how you believe you have learned about assessment during that time." The Round 2 interview asked teachers to describe the processes they had used in preparation for moderation and the sources of their practices. The Round 3 interview asked teachers to reflect on the changes in assessment made during the research period and the sources used. Questions were based on questionnaire responses to the items in Table 14 concerning the phases of transformative learning. Participants were asked about noteworthy events, times they had felt challenged and whether they had consciously evaluated the changes in assessment. Thus participants might identify the quality management processes as a source of learning, when the processes were not the direct focus of the questions.

#### **3.4.4.1.3 Teachers' assessment beliefs**

The interviews also collected data on teacher assessment knowledge and beliefs, frequently without requiring explicit questions. The Round 1 open question about the planning forums, for example, induced most participants to talk about their beliefs on the changes in SBA and the introduction of social moderation.<sup>27</sup> Where

---

<sup>27</sup> SACE Physics teachers were familiar with moderation as a statistical adjustment of school results based on comparisons of SBA marks and the external examination. Although teachers were aware that this was to be replaced by social moderation, details of the new form of moderation were not a focus of planning or clarifying forums. The term

this did not happen, questions were asked such as: “Can we look at your knowledge and beliefs about assessment, in particular looking at SBA. What are the advantages of including SBA in students’ final grades?” The Round 3 interviews asked participants to reflect on any changes in their assessment beliefs during the year. Scrutiny of questionnaire responses was used in preparing the questions for each interview. Thus, for example, questions on participant understanding of the purpose of the LAP were based on responses to the statement that the LAP was “merely a bureaucratic exercise”.

The Round 3 interview questioned differences of two or more between responses in Rounds 1 and 3, with differences of one considered insufficient evidence of change. Thus, for example, one teacher was asked: “In the questionnaire was a statement that some teachers will mark generously. Last year you were non-committal but this year you strongly disagreed. What has caused this change in opinion?” In this way data were sought on changes in assessment beliefs and practices over the research period.

#### **3.4.4.1.4 Teachers’ assessment practices**

Participants frequently revealed their assessment practices without specific questions being asked. Practices were described, for example, in responding to questions seeking clarification of questionnaire responses or when describing their experiences of quality management processes, such as attending a clarifying forum. Practices were also mentioned when responding to questions about assessment beliefs.

Although sometimes unnecessary, interview protocols included specific questions about assessment practices such as: “Can you describe any change you have made, or may make, in your assessment because of the clarifying forum? What was it in the clarifying forum that could cause this change?”

Assessment practices were a focus of the second round interviews, which asked participants about their preparations of materials for moderation. These interviews included questions such as:

---

*moderation* was not defined for the interviews. Participants were asked about the purposes of moderation and possible benefits and disadvantages of social moderation.

- How did you use the specific features in designing your Issues Investigation?
- How did you use the performance standards in arriving at the grade level for each student?

The interviews asking participants to critique their own assessment tasks (see 3.4.4.3) also sought information about participants' assessment practices.

#### **3.4.4.2 Verbalisation of assessment practices**

The quality assurance processes aimed to educate teachers in assessment activities, such as planning a program of assessment, designing assessment tasks, and using performance standards in grading student work (e.g., SACE Board of South Australia, 2010e).

The verbalisation exercises tried to identify whether participants had developed skills with LAPs and in using performance standards. Crisp (2008a) had used “verbal protocol analysis” with examination markers. The current research envisaged participants undertaking three exercises: reviewing a LAP, grading a piece of student work and critiquing another teacher's grade.

Verbalisation in this research varied from previous research in a number of ways. First, it involved SBA whereas previous research concerned examination markers (Crisp, 2008a, 2010b; Greatorix & Suto, 2006). Second, earlier research allowed participants to practise the protocol prior to data collection (Crisp, 2008a, 2010b; Greatorix & Suto, 2006). Practice was precluded from the current research because it might serve as a learning tool, influencing what was being measured. Previous verbalisation research had involved experienced assessors, presumably with established behaviours. Because the effect of practice was untested, it was excluded from the current protocol. It is accepted that the lack of practice might affect the reliability of the evidence. While verbalisation does not affect marker reliability (Crisp, 2008b), the effects of practice were unknown. Third, whereas earlier research sought to describe assessors' practices, the current research sought evidence of change in practices as evidence of learning.

The protocol<sup>28</sup> asked participants to “tell me what you are thinking and saying to yourself, as you work on these tasks” and recommended: “keep talking all the time as you work.” They were asked to include as much detail as possible, including changes of mind. In the current research responses were not continuous; participants were inclined to comment after completing a task rather than while undertaking it. The verbalisation hoped for direct evidence of teachers’ practices, distinct from later recollections as in the other interviews.

Because the current research was looking for evidence of change in assessment practices, the verbalisation activities were undertaken in Rounds 1 and 3. The sessions were recorded and transcribed. The planned analysis of the verbalisation activities are described later (see 3.9.5.2).

#### **3.4.4.3 *Own task critique***

In critiquing one of their own assessment tasks, participants were asked what had changed, why they had changed it as they had and what resources had been used.<sup>29</sup>

#### **3.4.4.4 *Summary of interviews***

Table 18 shows the data collected in each interview round: the focus of each interview, the topics covered and the research questions addressed.

---

<sup>28</sup> See Appendix N.

<sup>29</sup> The protocol is presented in Appendix O.

**Table 18: Data to be collected in each round of interviews**

Focus and topic	Round			Research Questions			
	1	2	3	1	2	3	4
<b>Assessment experiences</b>							
Experiences of quality management processes	✓	✓	✓	✓		✓	✓
Planning forum	✓			✓		✓	✓
LAP approval	✓			✓		✓	✓
LAP feedback	✓			✓		✓	✓
Clarifying forum presenter		✓		✓		✓	✓
Clarifying forum participant		✓		✓		✓	✓
Moderator			✓	✓		✓	✓
Moderation feedback			✓	✓		✓	✓
Learning about assessment	✓	✓	✓				✓
Teachers' assessment beliefs	✓		✓				✓
<b>Verbalisation</b>	✓		✓				✓
<b>Own task critique</b>	✓						✓

### 3.4.5 SACE Board officers focus groups

The fifth form of data collection involved focus groups of SACE Board officers, who had been involved in the quality management processes. The group format was selected to “stimulate talk from multiple perspectives” (Bogdan & Biklen, 2007, p. 109), anticipating that the group might provide an “opportunity for interaction ... and encouragement of ideas and insights” (Owens, Shute, & Slee, 2001, p. 24). In focus groups, some individuals may be reluctant to share “important experiences they have had because they are too embarrassed to share them in the group” (Bogdan & Biklen, 2007, p. 109). Thus, because teachers might be unknown to each other, and to maintain anonymity, focus groups were not employed for teachers. They were believed appropriate, however, for SACE Board officers. First, focus groups can be “advantageous ... when interviewees are similar to and cooperative with each other” (Creswell, 2008, p. 226), which was so; they were familiar with each other and with the interviewer. Second, the subject under discussion was not threatening—teachers’ learning opportunities, rather than the officers’ effectiveness, satisfying the “simplest test [of a focus group,] ... how acutely and easily the participants would discuss the topic” (Morgan, 1997, p. 17). Evidence of the focus groups’ success lies in the ease with which conversation flowed, with all officers contributing.

As with the teacher interviews, participating officers were asked to verify the accuracy of the transcripts, which were imported into NVivo for analysis.

The decision to use focus groups for SACE Board officers also resolved an ethical issue. Focussing the current research on Physics teachers had an implication for the anonymity of the one officer who delivered the quality management processes in that subject. Using focus groups for collecting data from SACE Board officers accessed others involved in the quality management processes in other subjects. This assisted in the anonymity of the officers and gave a broader picture of teachers' responses to the quality management processes.

The officer focus groups were timed to occur in conjunction with the first and third rounds of teacher questionnaires and interviews. Protocols<sup>30</sup> were developed listing questions to be addressed in the sessions; these were open-ended so that participants were free to frame their responses, with clarification sought where necessary. Each topic was introduced with a lead question which, in most cases, generated detailed responses; for example: I would like to hear your comments on teacher learning that was intended in the planning forums. The protocols included sub-questions which were used when detailed responses were not forthcoming.

Officers were asked to comment on teacher involvement and to describe the quality management processes and the opportunities they provided for teacher learning. They were asked what learning was anticipated for teachers, and what features of the processes led them to believe they provided opportunities for teacher learning. They were asked to describe any evidence that teacher learning had occurred. In this way, data contributed to all four research questions.

### **3.4.6 Data collection overview**

Table 19 shows the relationships between the forms of data collection and the four research questions.

---

<sup>30</sup> These are presented in Appendices P and Q.

**Table 19: Relationship between data and research questions**

	<b>RQ 1 teacher involvement</b>	<b>RQ 2 form of QM processes</b>	<b>RQ 3 content of QM processes</b>	<b>RQ 4 evidence for teacher learning</b>
Documents		✓	✓	
Observation		✓	✓	
Each round: Teacher questionnaires	✓		✓	✓
Longitudinal data: Teacher questionnaires				✓
Each round: Teacher interviews	✓	✓	✓	✓
Longitudinal data: Teacher interviews				✓
Officer focus groups	✓	✓	✓	✓

### 3.5 Delimitations

Delimitations refer to “drawing the boundaries around a study, ... showing clearly what is and is not included” (Punch, 2006, p. 69). They refer, therefore, to conscious decisions made in the planning stages to restrict the research.

The decision to recruit Physics teachers as participants was explained earlier (see 3.3). It is acknowledged that, by restricting the present research to teachers of one subject, generalisation to teachers of other subjects would be fraught.

The intention to recruit teachers using the SACE Board’s online forums might suggest greater participation by teachers with better online access but, as discussed earlier (see 3.4.3) this was not considered problematic and allowed participation by a geographically diverse group of teachers.

A decision was made to restrict evidence for teacher learning to assessment beliefs and practices in SBA, even though claims exist for social moderation providing professional development opportunities in teaching and learning (e.g., Crafter et al., 2006; Harlen, 2004a; Hipkins & Robertson, 2011; Matters, 2006). Another decision made was to restrict the personal data collected to gender and assessment beliefs and practices, even though teachers’ professional identities influence their learning and



behaviours (Adie, 2013; Bissaker, 2009; Day & Gu, 2007; Timperley, 2008; van den Berg, 2002).

### **3.6 Limitations**

Limitations are weaknesses “unavoidably present in a study's design” (Punch, 2006, p. 69). Although a range of opinions was collected and there was diversity in participant teaching experience, location and school sector, restricting the current research to the teachers of one subject, and the small number of participants, limited the validity of any generalisations.

The ethical requirement to obtain informed consent from participants meant that participation was based on “willingness to participate ... rather than on systematic sampling strategies” (J. P. Gall et al., 2005, p. 130). The use of the SACE Board online forum for Physics teachers avoided opportunity bias, “targeting groups likely to respond the way you wish” (J. Anderson & Poole, 2001, p. 26), but was not able to address response bias, in which “only motivated individuals within the groups targeted respond” (J. Anderson & Poole, 2001, p. 26).

Some of the participants had experience of writing a LAP and might have become familiar with SACE Board terminology and documents from teaching Stage 1 (Year 11) subjects in 2010. Opportunities for involvement in SACE Board quality management processes were, however, restricted to attending planning forums and having LAPs approved. Teachers did not participate in LAP approval, there were no clarifying forums and moderation of results was limited. Although teacher involvement was less than in 2011, Stage 1 experience might have informed teachers’ beliefs and practices prior to the current research.

It was recognised that teachers might become better informed about assessment in general, and social moderation in particular, through involvement in the current research. Swan (2006) reported questionnaires became more than an evaluation tool; they also became “a powerful way of provoking reflection and discussion among teachers” (p. 58). Thus participation in the current research might have similarly stimulated participants. It is impossible to ascribe any observed changes over the current research period to involvement in the quality management procedures; the changes might be attributable to research participation or other aspects of the

participants' lives.

There was a risk that the opportunity to review interview transcripts might be used to amend, not only wording, but opinions. Earlier research (B. Atherton, 2009) found that changes made by participants in similar research were editorial in nature and did not attempt to modify opinions expressed. In the current research, the very few amendments made to transcripts were always of a clarifying nature.

Variations in interview location, duration or timing might have generated differences in responses. All interviews were conducted at a time and location of the interviewee's choice. One interview was held at a university; the interviews for one teacher were conducted in the teacher's home, close to the school. All other interviews were undertaken at the teachers' schools either during, or immediately after, the school day. There were differences in the proximity of the interview dates to the corresponding SACE Board quality management process. Usually data collection occurred approximately three months after the relevant quality management process. Participants nominated interview times and the clarifying forum attended, generating variations in the time between process and interview. Variations were small, with one exception where the final interview was delayed because of teacher illness.

Rapley (2004) described how variations in the text and attitude of an interviewer cause differences in participant responses. As the current research progressed, some variation of questions occurred in order to pursue points raised by other participants. Variations also occurred because of an interviewee's comments: "The point is to follow the interviewee's talk, to follow up on and to work with them and not strictly delimit the talk to your predetermined agenda" (Rapley, 2004, p. 18).

Interviews are not impartial data collection but involve interactions resulting in "negotiated, contextually based results" (Fontana & Frey, 2000, p. 246). Familiarity with the interviewer might affect the interview process and my past roles with SSABSA and as a SACE Board officer meant that I was known to some of the teachers and all of the SACE Board officers who participated in the current research. Familiarity with their situations meant that the occasions when clarification was needed were few and I was able to empathise with them, so that elaboration was

forthcoming when sought. Twelve of the thirteen teachers undertook interviews in all rounds; the one who declined the final interview completed the final questionnaire but was not teaching Physics in 2012. The fact that so many participants completed all the scheduled interviews suggests that they were comfortable.

Participant willingness to “talk freely about their points of view” (Bogdan & Biklen, 1998, p. 95) was demonstrated when the participants disclosed aspects of their assessment processes about which they were uneasy. One, for example, was worried by “pre-testing” of students and another was concerned over “atomisation” of task instructions amounting to scaffolding. Holstein and Gubrium (1997) described the establishment of “climate for mutual disclosure” (p. 119) in which the interviewer is also willing to share feelings and opinions “to assure respondents that they can, in turn, share their own thoughts and feelings” (p. 119). In instances such as these, supportive comments enhanced the atmosphere of the interview.

The relative lengths of interviewee and interviewer contributions in transcripts can provide evidence of interview quality: “If the parts labelled [interviewee] are long and those designating the interviewer are short, you usually are looking at good, rich interview material” (Bogdan & Biklen, 1998, p. 99). To examine interview quality, analysis was undertaken of a random sample of four Round 1 interviews. Word count data showed that interviewees’ words comprised 87%, 69%, 89% and 74% of each interview. The mean of 80% suggests that the interviews represented “good, rich interview material”.

The interviewer was a 67-year-old male and the possibility was considered that, for the women interviewed, gender and age differences might have limited the interview effectiveness. The four interview analyses reported above included one female. The transcripts of the first interviews with the other three females showed that interviewees’ words comprised 87%, 81%, 75% and 82% of each interview: means were 81% for females and 77% for males, suggesting minimal gender effect.

Verbalisation has been shown not to interfere with the reliability of assessors’ judgements when used with experienced markers who have had practice in the protocol before its use (Crisp, 2008b). In the current research, the protocol was used without opportunity to practise it and with participants varying in expertise.

Consequently, findings based on verbalisation data were treated as tentative.

Another limitation of the current research related to the duration of the study. Loucks-Horsley et al. (2003) described three years being needed for teachers to fully implement changes in teaching practices, while Hipkins and Robertson (2011) described the New Zealand NCEA standards emerging over a number of assessment cycles. Ingvarson (1990) reported that, after the first year, a small majority of teachers had endorsed the moderation processes in Victoria, but that the number approving rose with more experience. That the current research was restricted to one assessment year is acknowledged as a limitation.

### **3.7 Ethical matters**

Ethics approval for the current research was granted by the Social & Behavioural Research Ethics Committee of Flinders University. The Committee approved the approach to the SACE Board of South Australia for permission to access teachers through the SACE Board online forums and SACE Board officers by email. Because the research did not involve children or vulnerable adults, no other permission was required.

The anonymity of participants has been protected by the use of pseudonyms. As mentioned earlier, the use of focus groups for data collection from SACE Board officers also served to protect the identity of officers. Reluctance to describe individual teachers was because combinations of teacher characteristics might facilitate identification. Participant descriptions have been presented holistically.

### **3.8 Evolution of the research**

When teachers were invited to participate in the first questionnaire, only eighteen teachers responded from a population in excess of 150. This limited response meant that the anticipated broad-brush picture from the questionnaires became unachievable; generalisation from the sample to the population would be invalid. Of the eighteen, however, thirteen agreed to be interviewed. Consequently, the decision was made to limit subsequent questionnaires to those teachers who had agreed to interviews and forego the information provided by the others.

During the focus groups conducted with SACE Board officers, mention was made of

exit surveys conducted in planning and clarifying forums. Access to these materials was granted by the SACE Board on the proviso that individual teachers or schools could not be identified. These reports (SACE Board of South Australia, 2010a, 2010b, 2011h) are cited in this thesis.

Prior to the final round of interviews, the SACE Board announced an evaluation of the first year of the new SACE (SACE Board of South Australia, 2012g). The final round of interviews was amended to include questions about suggestions the teachers might make and the changes they expected. Research is an “uncertain and messy business ... Particular lines of investigation often have to be abandoned ... while quite different possibilities emerge over the course of the investigation. The route followed is rarely a straight one” (Hammersley, 1997, p. 136) or, more poetically, “The best-laid schemes o’ mice an’ men gang aft agley” (Burns, 1785).

### **3.9 Preliminary treatment of the data and approaches to data analysis**

Although the absence of predetermined categories of analysis affect “the depth, openness, and detail of qualitative inquiry” (Patton, 2002, p. 14) the current research was undertaken believing it was impossible for a researcher to have “no a priori assumptions” (Grix, 2004, p. 114). Grix (2004) described deductive research as a strategy “in which theory informs research at the outset ... [in which data are] collected to confirm or falsify the hypotheses” (p. 113). It is accurate to say that the current research was informed by theory at the outset, since one incentive sought substantiation of statements that social moderation contributes to teacher professional development; however, the current research did not examine social moderation alone but included other SACE Board quality assurance processes. Furthermore, because the design was not experimental, the research was unable to “confirm or falsify” (Grix, 2004, p. 113) the literature claims of professional development for teacher involvement in social moderation. The research was broader than examining social moderation since data were also collected on other quality management processes, but was restricted to those of the SACE Board.

Although the research questions grew from claims made in the literature, it was exploratory rather than experimental. Thus it included elements of both deductive and inductive approaches to data analysis. While “deduction and induction can both

be used” (Punch, 2006, p. 24), “most qualitative research is both inductive and deductive” (Liamputtong & Ezzy, 2005, p. 57). Grix (2004) explained that “most research uses both induction and deduction, as there is a necessary interplay between ideas and evidence in each research process” (p. 114).

A deductive approach was used in writing research questions, designing questionnaires and creating protocols for the interviews and focus groups. While, inevitably, the research was guided by the researcher’s beliefs (Denzin & Lincoln, 2005), attempts were made to work inductively, to have findings “emerge out of the data” (Patton, 2002, p. 453). To improve the possibility of discovering the unexpected, I tried to “hold in abeyance anticipations, hunches, expectations, hypotheses, and so on” (Rennie, 2000, p. 485). Coding began with teacher interviews and the officers’ focus group from Round 1. Analyses of documents, observations and questionnaires were delayed to minimise their influence on the inductive coding of the interviews and focus groups.

Because the approach to analysis varied with the method by which the data were obtained, the treatment of the different data have been described separately.

### **3.9.1 SACE Board documents**

SACE Board documents will be described first because they were used in constructing the research questions, the questionnaires and the protocols for interviews and focus groups.

Document analysis was deductive: each was analysed “according to an existing framework” (Patton, 2002, p. 453). Documents were read for instances concerning the form and content of the quality management processes (Research Questions 2 and 3) as the basis of analysis (M. D. Gall, Gall, & Borg, 2007).

Thus evidence was sought that the quality management processes might display features of effective professional development, effective teacher learning, deep learning and adult learning. For example, the documents were examined for “content focus, active learning, coherence, duration, and collective participation ” (Desimone, 2009, p. 183); teachers “working with their colleagues in professional learning communities” (Wei et al., 2009, p. 9); opportunities for discussion (Thompson &

Zeuli, 1999) and teachers undertaking analysis of student work (Wei et al., 2009, p. 12). Evidence was also sought that teachers were aware of “why the change is desired ... and [had] clear goals of desired outcomes” (Bissaker, 2009, p. 13).

### **3.9.2 Observations of quality management processes**

Observations of the SACE Board quality management processes included a planning forum, LAP approval, preparation for clarifying forums and attending clarifying forums. As described earlier, field notes were made and comments added both during the observation and later.

Analysis of the field notes, like that of the documents, was deductive because they were read with Research Questions 2 and 3 in mind. Evidence was sought that the quality management processes might display features of effective professional development, effective teacher learning, deep learning and adult learning.

### **3.9.3 Treatment of missing data from questionnaires**

Questionnaires included both multiple-choice and open-ended questions and dealt with assessment beliefs, assessment practices, school support, experience of SACE Board activities, personal information and learning about assessment.

There were some instances of data omission. Although there are “no firm guidelines for how much missing data can be tolerated” (Tabachnick & Fidell, 1996, p. 60), it is believed that the missing data were not significant. First, there were few omissions. The first questionnaire had seven unanswered items from almost 1100. The second and third questionnaires had fewer omissions, with no participant exceeding one missing result. Coverage of the items was not hindered by the omissions as they were for different items, with the exception that two teachers failed to respond to the item: I believe that my next learning and assessment plan will be approved without amendment. The manner in which omissions were tackled is discussed below.

The literature describes a number of processes for dealing with missing data. With only thirteen participants, the option of dropping questionnaires with missing data (De Vaus, 2002; Tabachnick & Fidell, 1996) was not considered. A second option is to estimate missing values (De Vaus, 2002; Tabachnick & Fidell, 1996). One estimation method involves replacing a missing value with “a value from a well-

educated guess” (Tabachnick & Fidell, 1996, p. 63). A second form of estimation, uses the mean as “the best guess” (Tabachnick & Fidell, 1996, p. 63). A third solution can be applied to variables calculated from a number of items—to calculate the score using only the completed items (De Vaus, 2002).

The current research used a number of these options. The corresponding interview transcript was read to identify a “well-educated guess” (Tabachnick & Fidell, 1996, p. 63), which could be used. One example involved a missed response for the item concerning a LAP being merely bureaucratic. In the interview, the participant had said that the planning forums had stimulated his insistence on the early submission of his school’s LAPs. Consequently, it was decided that he saw the LAP as more than a bureaucratic exercise and his response was entered as *Disagree*. Another participant had not completed the item: External assessment is objective. However, responses to other items showed he believed that school based assessment was open to manipulation and that external assessment is the fairest form of assessment. In the interview these responses were discussed and he commented that, in external assessment, “Everybody’s on a level playing field.” Consequently, the missing response was entered as *Agree*.

The failure of two teacher-participants to respond to beliefs about future LAP approvals was similarly addressed: interview responses suggested confidence that future plans would be approved. Although the group mean was never used for a missing value, the group median was used in a few cases. The mean was not appropriate since the group mean was rarely an integer and so did not coincide with a point on the Likert scale used. The group median was used provided two conditions were fulfilled: the item showed little response variation and the participant with missing data had group-median responses for similar items. Other unanswered items were left incomplete. Where the item was used for calculating a variable, this was the equivalent of using the participant’s mean of the other items used in the calculation.

#### **3.9.4 Coding of interviews and focus groups**

Although an interview protocol was used, individual interviews varied greatly because questionnaire responses were used to personalise each interview and open questions led interviews in different directions. In building “a systematic account of



what has been observed and recorded” (Ezzy, 2002, p. 86), transcript analysis codes were applied to “segments, or chunks, of data” (Liamputtong & Ezzy, 2005, p. 331), so bringing “meaning to information” (Rossman and Rallis cited in Creswell, 2009, p. 186).

In grounded theory, a researcher generates theory by examining the data (Ezzy, 2002; Liamputtong & Ezzy, 2005; Patton, 2002), particularly useful “when existing theories do not address your problem or the participants that you plan to study” (Creswell, 2008, p. 432). The current research examined the quality management processes through the lenses of a number of theories: teacher learning, professional development and adult learning. It did, however, “hold in abeyance ... anticipations, hunches, expectations, hypotheses ... about the phenomenon of interest” (Rennie, 2000, p. 485), remaining open “to all possible theoretical directions” (Charmaz, 2006, p. 46). Consequently, coding of interviews and focus groups was undertaken prior to analysis of documents and observation records to minimise their influencing the coding. Likewise, reading of the First Year Evaluation (SACE Board of South Australia, 2012b) was postponed until data analysis had been completed.

Although a researcher attempts to “capture the empirical reality” (Charmaz, 2006, p. 47), codes inevitably show the researcher’s view because the researcher chooses the words constituting the codes (Charmaz, 2006). Researchers define what is significant data and describe what they “think is happening” (Charmaz, 2006, p. 47). Doubtless my pre-conceptions and assumptions influenced the analysis in the current research, despite attempts to minimise their influence.

Charmaz (2006) described coding as consisting of at least two phases: an initial phase in which the researcher remains “open to all possible theoretical directions” (p. 46) and a second phase in which the codes are organised. The second phase requires decisions that make the most sense of the data. Creswell (2008) described the second phase as “making connections between categories” (p. 460) by examining the codes for “overlap and redundancy, and [collapsing] these codes into broad themes” (p. 251). These two phases are commonly called *open* and *axial coding* (e.g., Corbin & Strauss, 2008; Creswell, 2008; Ezzy, 2002; Liamputtong & Ezzy, 2005). Many grounded theorists identify a third phase, *selective* (e.g., Creswell, 2008; Ezzy, 2002; Liamputtong & Ezzy, 2005) or *theoretical* (Charmaz, 2006) coding, in which theory

is developed that “provides an abstract explanation for the process being studied” (Creswell, 2008, p. 437).

Corbin and Strauss (2008) described the distinction between the phases as “artificial” (p. 198), although it does indicate that the data are broken apart and then put “back together again” (p. 198). In the current research, coding was an iterative process in which “the steps [did] not occur in a linear fashion” (Bissaker, 2009, p. 61), although it could be considered to involve three phases. The first two approximated to open and axial coding. These terms are used in grounded theory, whereby theory emerge from the data. Because the current research used existing theory to scrutinise the phenomena studied, grounded theory terms were considered inappropriate and the three phases were titled *exploration*, *consolidation* and *theorisation*.

#### **3.9.4.1 Coding phase 1: exploration**

Open coding has been described as “exploratory” (Ezzy, 2002, p. 87) or the “first run” (Liamputtong & Ezzy, 2005, p. 268) at coding data. This description fits the exploration phase of coding in the current research, involving an inductive approach. No pre-defined codes were used, although interview questions might have been used in this way and it might have been possible to anticipate codes for teachers’ beliefs, practices and feelings.

Code development entailed line-by-line scrutiny of the transcripts of interviews and focus groups for segments that might have something to say, such as positive and negative comments about the processes, events participants recalled, actions they undertook, reasoning behind actions, how they were feeling and clarification of comments made in questionnaires. In this process comparisons were made between segments (Liamputtong & Ezzy, 2005) and, where similarities were evident, existing codes were assigned. New codes were generated when similarities with existing codes were not evident. A number of techniques were used to assign codes, short names that “summarize” (Charmaz, 2006, p. 43) each segment. Although organisation of codes belongs to the consolidation phase, some grouping and revision were undertaken during the initial coding exploration. For example, Russell said the timing of the planning forums was “about right”. When others said they were “too early” they were both coded as *Timing*.

*In vivo* coding uses “the exact words of participants” (Creswell, 2008, p. 440); possibly the words used are “strikingly original” (Booth, Colomb, & Williams, 2008, p. 97). One teacher, for example, referred to “clarifying forums that don’t clarify” while a second worried about doing his students a “disservice”. A third described the holistic approach to grading as “humming”, whereby he would “read the whole thing, hum a little and produce a grade or a mark”. In other cases participants used expressions “that flag condensed but significant meanings” (Charmaz, 2006, p. 55) that would be understood by other teachers. Examples in the current research included “playing the game”, “it’s all in my head”, “policy on the run” and “looking good for moderation”.

*In vivo* codes used are a consequence of the order of coding. For example, an early transcript generated the code “airy fairy” to describe the performance standards. An interviewee coded later used the term “wishy washy”, which was coded as “airy fairy”. Had these interviews been coded in reverse order, the coding used would have been “wishy washy”. Thus, with *in vivo* coding, the order of coding may affect the term used for a code.

A second form of assigning codes uses terms from the literature. Thus, for example, one participant said: “I adapted it a bit to make sense in my head.” This was coded as *making sense* (Moon, 2000). Other passages were coded as *making meaning* or *working with meaning*, reflecting Moon’s stages of learning. Participants used a number of terms for their beliefs about the focus of the changes to the SACE, including “education-for-all”, “political” and “retention”. These were coded using the title of the SACE Review (Crafter et al., 2006), *success for all*.

As the number of codes grew, a decision was made to begin consolidation after coding a sample of interviews. The thirteen participants included six with school leadership roles: three faculty heads and three curriculum leaders. Six participants were SACE Board moderators in 2011 and four did not moderate in 2011 and did not hold school leadership roles. Exploratory coding was completed for all three interviews of four participants: one faculty head, one curriculum leader, one moderator and one teacher with no SACE Board panel involvement. From each category the participant used was the one who had generated the greatest range of codes in Round 1 interviews. Because *verbalisation of assessment practices* and *own*

*task critique* might generate different codes from the initial interviews, it was decided to code two of each of these using the two participants with the greatest range of codes in Round 1. Because SACE Board officers represented a different viewpoint from teachers, it was decided to code both focus group transcripts before consolidation.

The extensive exploration prior to consolidation minimised the need to generate new codes for subsequent transcripts.

#### **3.9.4.2 Coding phase 2: consolidation**

A grounded theorist, in the second phase of coding, selects one code and then “relates other categories to it” (Creswell, 2008, p. 434). Consolidation in the current research better fits other descriptions: examining relationships (Bissaker, 2009; Corbin & Strauss, 2008) or making connections between codes (Creswell, 2008). Thus organisation is generated among the codes. The first phase breaks down data into labelled segments (Creswell, 2008), while the second puts “data back together in new ways” (Strauss & Corbin cited in Liamputtong & Ezzy, 2005, p. 269). In consolidation, codes were examined for overlap and redundancy (Creswell, 2008). Further grouping was also undertaken, extending that undertaken in exploration.

During consolidation merging and renaming of codes occurred. Some initial codes were retained, while codes expressing the same concept were merged. Some participants, for example, had spoken of “disadvantaging” their students rather than “doing students a disservice”; all were coded as “disservice” even though this term was not used unanimously. On the other hand, although the picture of a teacher “humming” was appealing, the code became *holistic* since this better reflected the description given by others.

In questioning the effectiveness of moderation, some participants had referred to the *lack of time* available; others suggested it would be restricted to *tick and flick* procedures and others that it would be *time-consuming*—all had been coded differently. After examining the segments all were coded as *time-consuming*. A second example arose in participants’ different expectations of learning from moderation, which had been coded as *see other teachers’ work*, *see work from many schools*, *teacher discussion* and *moderator training*. While the different codes were

retained, they were grouped under *moderation - teacher learning*.

Some codes evolved during the coding process. A description of SACE Board responses to concerns raised by teachers had been coded *lack of response*.

Interviews coded later included the following segments:

- We were listened to, but I interpreted that as they wanted to be seen to be listening.
- I think they were listened to but I'm not sure that they were actioned upon.
- I think they were listening, I don't know whether they were valuing.
- They were listened to, yes. But it was really, "This is the new system ..."

These were considered dismissive of teachers rather than a lack of response.

Consequently, the code was changed from *lack of response* to *response to teachers*.

The current research was not considered to be grounded theory because, although codes arose from the data, there was not an absence of preconceptions; a number of theories were applied in examining the quality management processes and it was possible to group the codes in different ways. Table 20 shows several strategies that were considered for consolidation of codes.

**Table 20: Alternative strategies for consolidating codes**

	<b>Strategy 1</b>	<b>Strategy 2</b>	<b>Strategy 3</b>
Level 1	Quality management processes	Research questions	Research questions
Level 2	Research questions	Quality management processes	Sub-questions
Level 3	Sub-questions	Sub-questions	Quality management processes

The first two strategies involved much repetition. Codes related to reasons for involvement, for example, would be repeated for each of the processes since the same reasons were offered for participation in each process. Consequently, the third strategy was adopted. The research questions became the primary organiser, with the sub-questions the second level, as shown in Table 21 for the second research question concerning the form of the quality management processes.

**Table 21: Coding organisation based on research questions and sub-questions**

Level	Code
1	Form
2	Professional development activities
3	Active learning
3	Coherence
3	Duration
3	Collective participation
2	Effective teacher learning
3	Embedded in teachers' work
3	Moderation during year
3	Challenging prevailing discourses
3	Active learning
3	Active learning
4	Dissonance
4	Resolving dissonance
2	Adult learning
3	Planning input
3	Relevance
3	Congruence
3	Active learning
3	Materials

Theorisation was undertaken after exploration and consolidation had been completed for the sub-set of transcripts described above. The remaining transcripts were coded following the first round of theorisation.

This organisation was effective for the data on teacher involvement and the form and content of the quality management processes but was revised for the fourth research question concerning teacher learning. Reading had continued and one article described teacher learning as occurring in individual singular episodes (Evans, 2014). The article referred to the interconnected model of teacher change (D. J. Clarke & Hollingsworth, 2002) and the concept of change sequences (see 2.2) which had not been considered previously. It was therefore decided to revise the fourth research question, related to teacher learning, and recode all teacher transcripts for teacher learning. The organisation, similar to that shown in Table 21, included new

codes for the revised picture of teacher learning.<sup>31</sup>

It was also decided that the learning journey of each teacher would be more apparent if the transcripts for each teacher were read consecutively. In exploration, transcripts had been coded in the order of their recording: first round interviews had been followed by the second round and third rounds. Similarities and differences between participants helped code creation during exploration and assisted consolidation during exploration. In the re-reading, transcripts for each participant were read in sequence: all of Adam's transcripts were read and coded consecutively, followed by the others' in alphabetical order.

### **3.9.4.3 Coding phase 3: theorisation**

In the third phase each code was scrutinised and a decision made to retain or reject it based upon whether it related to one of the quality management processes and to one of the research questions. A code satisfying both criteria was retained and categorised accordingly, "with the research questions used as a lens" (Connolly et al., 2012, p. 601). Other codes were rejected, although a few were retained that appeared "interesting" ... in order to facilitate new discoveries" (Liamputtong & Ezzy, 2005, p. 281). Some of these were eventually used. One example was the coding for suggestions that some moderation samples be judged by a second pair, unaware that the sample had been previously moderated. These were initially coded as *blind test* since this term was used by one participant. This was later renamed *quality review*, as an example of this component of quality management. Others that were retained included *critique of research* and *learning from research participation*.

### **3.9.5 Evidence of teacher change**

Evidence of learning can be obtained by asking the learner to reflect on the learning and by comparing the learner's knowledge and skills before and after the learning event. Both were used in the current research.

#### **3.9.5.1 Participant reflection**

Participant reflection on their learning had been used in previous research (e.g.,

---

<sup>31</sup> See Appendix S.

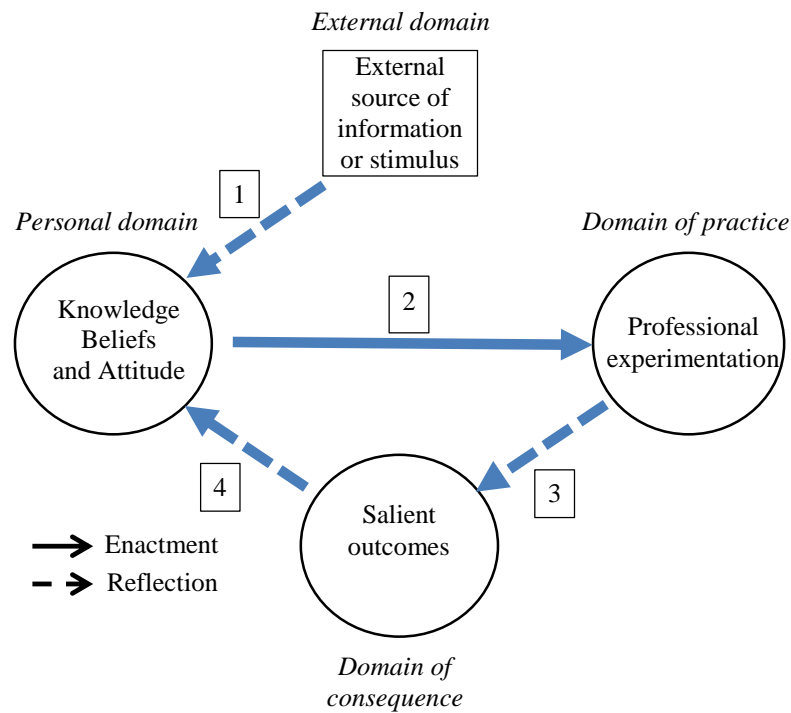
Limbrick & Knight, 2005; L. Reid, 2007) into the educative value of moderation. In the current research, questions were asked about evidence that teachers had learned from the quality management processes. Responses were coded to describe what had changed. One teacher, for example, stated: “I think my belief about what SACE Board calls ‘good assessment’ may have changed.” This was coded as a change in assessment beliefs. Another example of a change in beliefs was provided by the teacher who explained how, previously, good assessment had meant “getting the kids in the right order” because statistical moderation would make any upward or downward adjustments; this was contrasted with the new system where “you’ve got to make sure that everything matches the way the moderator wants it to”. Very common were changes made in task design, such as assigning specific features to test questions, implementing ideas from other teachers and reflecting on their own tasks.

All participants described change sequences (D. J. Clarke & Hollingsworth, 2002), episodes of micro-level professional development (Evans, 2014), most frequently in describing the development of an assessment task—own task critique—and in describing the preparation of materials for moderation. Initially these passages had been coded as *teacher learning*. In theorisation, these were analysed using the interconnected model of teacher learning (D. J. Clarke & Hollingsworth, 2002). For example, Adam explained his changed practice and beliefs regarding LAPs:

I guess I thought they [LAPs] were useful because from a personal point of view to just look at the way of planning Stage 2 subjects. ... going into Stage 2 even though assessment plans have been encouraged ... they didn’t have to be submitted. Often they were things that were done in-house and I guess we’ve got four teachers on staff who’ve been teaching Stage 2 Physics for a number of years. And I guess whilst that’s a great pool of experience to have, at times it becomes a bit like that knowledge is often internalised and not documented very well and it’s “Oh, we’ll do this next and we’ll do this and that’s how we’ll do this” and those kind of things. ... And so I relied a lot on talking to colleagues, saying “Where are you up to now? What are you doing?” Whereas by going ahead and having a formal assessment plan gave it a bit more structure. So if something were to happen and someone was to fall off the perch or something like that, it was all documented. So there is that benefit, ... it gave that opportunity to document what people were doing and why they were structuring the courses the way they did. So I thought that was one of the benefits of that kind of approval and planning.

This was analysed and mapped onto the interconnected model (Figure 9).





**Figure 9: Adam's learning regarding LAPs explained using the interconnected model of teacher learning**

In this sequence the external domain stimulus was the requirement to have the LAP approved by the SACE Board. Reflection (Arrow 1) led to a changed belief: the need for approval “gave that opportunity to document what people were doing and why they were structuring the courses the way they did”. Enactment (Arrow 2) was the development of a new LAP that was submitted and approved. Adam valued the outcome (Arrow 3) since he saw that the LAP served a purpose: “if something were to happen and someone was to fall off the perch or something like that, it was all documented. So there is that benefit”. He later added that “it has been good that it has been articulated in a format that other people can come in and look at it and see what is going on”. Reflection (Arrow 4) led to a changed belief: the “approval and planning” of a LAP was beneficial because it had “given it a bit more structure”.

Similar mapping was undertaken with each instance of teacher learning identified, with the findings discussed in the chapter on teacher learning (see 7.3). The transcripts were also analysed to determine whether participants considered the changes to be improvements (a “better way” as described in Evans, 2014).

### **3.9.5.2 Longitudinal data**

The current research sought to identify changes in teachers' beliefs and practices during the research period. It was anticipated that participants with greater involvement might show greater changes than teachers with less involvement. Difference between Rounds 1 and 3 were sought in a number of ways.

Comparison was undertaken of questionnaire responses from Rounds 1 and 3 prior to the final interviews. Any difference of one in the 5-point scale was considered to be insignificant as such a difference might easily arise were a teacher to complete the questionnaire twice within a few minutes. A difference of two (e.g., *strongly disagree* to *neutral*) or more was considered as suggestive of change. When this was observed, the participant was asked about it in the subsequent interview. In some cases the change was confirmed, in others the different responses were explained by the teacher answering the item differently because different aspects of the assessment cycle were relevant on the two occasions. In a few instances the teacher expressed surprise and believed a mistake had been made in the questionnaire. When this occurred the questionnaire response was amended. Questionnaire and interview data from Rounds 1 and 3 were analysed for changes in participants' beliefs concerning SBA, performance standards and social moderation.

A second method involved transcript analysis for use of SACE Board terminology and looked for evidence of change over the period of the research; an increase in correct usage, or a decrease in incorrect usage, of the terminology might be taken as learning.

A third procedure examined grades assigned to student tasks in the verbalisation activities. Had consistency, inter-rater reliability, improved over the year and had greater consistency developed between participants who had served as moderators?

An intended fourth approach also compared the verbalisations of Rounds 1 and 3. It was intended to identify cognitive strategies used in marking: *matching*, *scanning*, *evaluating*, *scrutinising*, and *no response* (Greatorex & Suto, 2006). It was found, however, that this was impossible as participants had not described their mental processing in sufficient detail. This may have been because of the nature of the task chosen; it may have been inappropriate for the strategies described by Greatorex and

Suto (2006). Alternatively, the cause might lie in the decision to not allow practice.

In a fifth approach, interview transcripts were analysed for evidence of participant confidence. Data were scrutinised for evidence that teachers with greater involvement in the quality management processes gained more in confidence than those with less.

A sixth approach examined participants' attitudes. The first and last interviews were scrutinised for positive and negative comments to find whether there had been any change over the research period.

Therefore many strategies were employed seeking evidence of change in participants over the research period. Data were examined for evidence of changes in beliefs, practices and attitudes.

### **3.10 Summary**

This chapter has described the longitudinal mixed methods design undertaken during 2011, the first year in which performance standards and social moderation were used in all high-stakes subjects in the SACE. The next chapter analyses the participants and their involvement in the quality management processes.

## **CHAPTER 4: RESULTS: TEACHER INVOLVEMENT**

In the next four chapters, the quality management processes will be examined from the viewpoints of the four research questions. This chapter begins with a description of the participants and an analysis of their representativeness of SACE Physics teachers. The chapter continues by examining teacher involvement in the 2011 quality management processes and whether it represented an increase over previous years'. Reasons for involvement or non-involvement are then explored, followed by an examination of teacher characteristics and the possible association of these characteristics with involvement. Participation in communities of practice and use of exemplars are then studied. The chapter concludes with an overview of teacher involvement in the quality management processes and the implications for teacher learning.

### **4.1 Participants**

Participants comprised teachers and SACE Board officers. Teacher data are examined and analysis presented of the extent to which they were representative of SACE Physics teachers. This is followed by a description of the participating SACE Board officers.

#### **4.1.1 Teachers**

Thirteen SACE Physics teachers agreed to be interviewed and became the participants for the current research. Five had non-SACE experience overseas, interstate and with the International Baccalaureate. Six were teaching other SACE subjects: sciences, mathematics, Information Technology and the Research Project. Three were faculty leaders and three held other school leadership positions. Consequently, participants reflected on the changes from many different perspectives.

For ethical reasons, participants will not be described individually since certain combinations of characteristics may allow identification. A small sample is unlikely to be representative of the population of SACE Physics teachers, but the data are examined for the extent to which they were typical.

#### 4.1.1.1 *Personal characteristics*

Four (31%) of the teachers were women. This over-represents women, given that 13% of SACE Physics teachers in 2012 were women (personal communication with SACE Board officer, 27 May, 2013).

#### 4.1.1.2 *School characteristics*

Data were collected on locations and sectors of participants' schools.

Eleven (85%) of the thirteen teachers were in metropolitan Adelaide. In 2011 71% of the schools with Physics classes (103 of 145) and 77% of Physics classes (139 of 181) were in metropolitan schools. Metropolitan schools were over-represented in the sample. In 2007, 82% of SACE assessors were from metropolitan schools (Kilvert & Mercurio, 2007). Therefore participants were in proportion to the population of Physics teachers.

One of the participants taught in a Catholic school, four in Government schools, and eight in Independent schools. These proportions are shown in Table 22 and compared to the sector representations in SACE Physics classes in 2011 (SACE Board, personal communication 7 June 2012) and the proportions of SACE Board assessors in 2005 (Kilvert & Mercurio, 2007).

**Table 22: Percentages of teachers in educational sectors**

	<b>Catholic</b>	<b>Government</b>	<b>Independent</b>
Participants	7	29	64
Physics classes	20	49	31
Assessors	20	53	27

Thus Independent schools were over-represented in the participants.

#### 4.1.1.3 *Professional characteristics*

Professional characteristics comprised participants' tertiary Physics background, the years they had been teaching, how long they had been teaching Year 12 Physics and prior involvement in SACE assessment panels.

Two participants had one year of tertiary Physics study and two had two years. Five had a major in Physics while four had post-graduate Physics qualifications. Nine of

thirteen, 69%, is slightly higher than the 58% of science teachers having a major in their teaching subject (Panizzon, Westwell, & Elliott, 2009, p. 12). The small sample size renders any comparison of these proportions invalid.

The length of the teaching experience of participants ranged from one to 38 years, with a mean of 22 years and a median of 20. Because data were not available, contact was made with six leaders of SACE Board Physics panels, who would be aware of the profile of SACE Physics teachers. Three suggested that the sample was, in regard to teaching career, probably representative of Physics teachers in the state. Two suspected the sample was possibly younger, while one suggested the sample was older. Panizzon, Westwell, and Elliott (2009) presented data ages, rather than careers, of South Australian science teachers. Their age profile (p. 8) yields a mean of approximately 45 years, suggesting that the average science teacher has been teaching for slightly more than twenty years. It appears that the sample might have been representative of South Australian Physics teachers as far as teaching career is concerned.

Two participants were teaching Year 12 Physics for the first time, while the most experienced teacher for the 34th time. The mean was 14 years and the median 12. There appeared to be no data that can be used to consider the representativeness of the sample.

Ten participants had previous experience of SACE Board involvement: eight (61%) as markers or moderators and six (46%) in various roles, including Chief Assessor, examination setter, curriculum writer or subject advisory panel member. Thus participation in SACE Board assessment panels was over-represented in the sample since, in 2005, 11% of Year 12 teachers were involved in “moderation and external assessment processes” (Kilvert & Mercurio, 2007, p. 7). The six participants (46%) who served on SACE Board assessment panels in 2011 also exceeded the participation reported by Kilvert and Mercurio.

#### ***4.1.1.4 Representativeness of the sample***

It appears that the participants were not representative of SACE Physics teachers. While gender, career duration and Physics teaching experience might be typical, three factors indicate that the sample was not representative: the high proportion of

teachers from Independent schools, the high proportion with previous SACE involvement, and the proportion serving on panels in 2011.

It was not surprising that teachers willing to be involved in panel commitments outside their teaching roles also volunteered to participate in a research project. The research suffered from a *selection-effects problem* (Donmoyer & Galloway, 2010).

#### **4.1.2 SACE Board officers**

SACE Board officers participated in a different way from the teachers. Teachers, as well as providing data, were the subjects of the research: evidence was collected on their participation, beliefs and practices. SACE Board officer involvement was restricted to providing evidence. Thus, while officers were treated as participants as far as recruitment and ethics were concerned, they served more as witnesses. Nevertheless, some profile of the officers is desirable.

The two focus groups were conducted approximately twelve months apart, in conjunction with the first and third rounds of teacher interviews. Six officers attended the first and five the second. Three attended both. Between the two meetings, two of the original participating officers had left the SACE Board and one was unavailable for the second meeting. Rather than proceed with three, an invitation was issued and two newly appointed officers participated in the Round 3 focus group.

The officers varied in experience from newly appointed to some with over a decade of SACE Board employment. The subjects they worked with included arts, sciences, mathematics and humanities.

This chapter now examines the data for evidence related to the first research question concerning teacher involvement in the SACE Board's quality management processes.

## **4.2 Teacher involvement in the quality management processes**

The SACE Board identified a number of processes that were part of their quality management cycle for school-based assessment (SBA). The use of exemplars and membership of communities of practice were included in the current research as

explained earlier. This section examines teacher involvement in the Board-specified processes in 2011 and whether this represented an increase over that of previous years. Teacher use of exemplars and involvement in communities of practice are then examined.

#### **4.2.1 Teacher involvement in planning forums**

Planning forums were conducted in the first half of 2010, at least six months before the changes were implemented in 2011. Data included questionnaires and interviews, as well as SACE Board exit surveys and the focus groups.

Nine participants attended at least one planning forum. Two did not attend because, at the time of the forums, they were unaware they would be teaching Physics in 2011. Another was informed of forum activities by the other two Physics teachers in the school, who had attended a planning forum. The final participant said that the planning forums had “slipped under the radar” at his school. Subsequently, SACE Board officers conducted a series of similar meetings at the school. Thus, it is considered that all who knew they would be teaching SACE Physics in 2011 had access to planning forum information.

SACE Board officers estimated attendances between 30% and 70% of the teachers of different subjects. They mentioned attendance difficulties for some country teachers because replacement teachers were unavailable, even though funds were provided. Timing of the forums was another difficulty mentioned. One officer said the teachers “weren’t ready” when the forums were presented: “By the time [the teachers] were ready to start learning about the processes, we had already finished all of the planning forums.” Conversely, one participant appreciated their timing: “If they had been any later I would have had a lot of trouble getting the staff ready”. Teachers who missed the planning forums were given online access to the materials used in the workshops.

The evaluation survey of the planning forums (SACE Board of South Australia, 2010a, 2010b) was completed by 101 Physics teachers. This constituted approximately 60% of the Physics teachers, agreeing with the participation rate suggested by the SACE Board officers. Because planning forums had not existed previously, comparisons of attendance with previous years is redundant.



#### **4.2.2 Teacher involvement in learning and assessment plan (LAP) approval and clarifying support panels**

LAP approval panels were observed in three subjects. Panels for subjects with large enrolments (e.g., Physics) consisted of approximately eight teachers. In 2011 the SACE Board appointed three supervisors in each subject: School Assessment, External Assessment and Materials (SACE Board of South Australia, 2010m). The three subject supervisors were members of LAP approval panels, supplemented by other teachers in subjects with large enrolments.

Two participants were supervisors and were involved in LAP approval, clarifying support and moderation, in addition to attending planning and clarifying forums as subject teachers. Thus supervisors experienced greater involvement in the quality management processes than other teachers.

#### **4.2.3 Teacher involvement in clarifying forums**

Questionnaire data on participant attendance at clarifying forums supplemented SACE Board data from an annual report and a focus group. During 2011 approximately 3000 teachers from the metropolitan area, and 1200 non-metropolitan teachers attended clarifying forums (SACE Board of South Australia, 2011b), but no subject breakdown of attendance was reported. A SACE Board officer (private correspondence) reported that 149 Physics teachers attended clarifying forums, confirming that clarifying forum attendance had exceeded that of planning forums, representing approximately 90% of Physics teachers. All thirteen participants had attended clarifying forums.

#### **4.2.4 Teacher involvement in moderation**

There were approximately 25 Physics moderators in 2011, including the three supervisors. Six participants served as moderators in 2011. Although this was an increase for Physics, incorporating social moderation for the first time, in subjects with previous social moderation (e.g., Chemistry), there was no increase in teacher involvement; moderation panels were of similar size to previous years.

#### **4.2.5 Increased teacher involvement?**

Numerically, teacher involvement in Physics quality management processes had increased over previous years. Closer scrutiny is necessary, however, because 2011

---

was the first year in which the processes existed for SBA. The following section scrutinises teacher involvement more realistically.

The data showed that 2011 involvement in the quality management processes created three groups of teachers with differing levels of involvement: supervisors, moderators and others.<sup>32</sup> Supervisors were restricted to three (approximately 2% of Physics teachers). The moderators, including the supervisors, represented approximately 17% of Physics teachers. For the majority, involvement was restricted to attending planning and clarifying forums.

All six moderator-participants of 2011 had previous SACE assessment panel experience. Five were examination markers, as well as moderators, in 2011. SACE Board officers reported that in subjects with a history of social moderation, a small number of new moderators was appointed, but there was no increase in the number of moderators; the new appointments appeared part of the “generative process” (Lave & Wenger, 1991, p. 57) of any community of practice. The widespread adoption of social moderation apparently had little impact on increasing teacher involvement in assessment panels. The increased teacher involvement anticipated by the SACE Review (Crafter et al., 2006) did not materialise. The reasons for this appeared to lie in the social moderation model adopted for the new SACE; this differed from that of Queensland, which had informed the Review Panel’s expectations. The differences between the social moderation models of Queensland and the new SACE are examined later (see 8.2).

SACE Board documents were examined to establish whether planning and clarifying forums were to be considered part of the ongoing cycle of quality assurance (SACE Board of South Australia, 2010f) or whether they were part of the implementation of the new SACE. Planning was identified as one phase in the SACE Board quality management cycle (SACE Board of South Australia, 2010f) and the first mention of planning and clarifying forums appeared in December 2009 (SACE Board of South Australia, 2010i), in which the SACE Board undertook to conduct planning and clarifying forums for teachers. Schools were expected to release teachers to

---

<sup>32</sup> To avoid cumbersome terms, these will subsequently be referred to as *supervisors*, *moderators* and *non-moderators* where the role is relevant.

---

participate in these forums.

The high attendance at the 2010 planning and the 2011 clarifying forums proved atypical. While funding had been provided for teachers to attend these forums, it was subsequently restricted to “teachers new to a subject without subject expertise in the school, for schools offering subjects for the first time, and for schools that have experienced significant moderation shifts in previous moderation” (SACE Board of South Australia, 2010i, p. 5), although other teachers might attend. In 2012 there was one meeting, for new Physics teachers, compared to twelve planning forums in 2010. In 2012 there were three Physics clarifying forums compared to twelve in 2011. It was estimated that 10–20 teachers attended planning forums after the first year, and 30 the clarifying forums. Most teachers will have attended one planning forum and one clarifying forum over their teaching careers. Apart from the two supervisors, no participant attended a planning or a clarifying forum in 2012. Officers explained how “a review of funding” necessitated a revision of what could be provided for teachers. It was concluded that these forums served as an induction to changes in the SACE, rather than part of an ongoing quality management cycle.

In summary, three SACE Physics teachers experienced extensive involvement in SACE Board activities due to their supervisory roles; moderation did not appear to increase the pool of teachers involved in SACE Board assessment panels. It was considered, therefore, that the quality management processes associated with SBA had not generated the increased participation recommended in the SACE Review (Crafter et al., 2006). Why had the new SACE continued placing assessment responsibility “in the hands of a small group of designated moderators” (Crafter et al., 2006, p. 130)?

Involvement in the quality management processes divided the research participants, as it had teachers, into three groups: two supervisors, four moderators and seven non-moderators. The 15% proportion of supervisors among the participants exceeded the proportion of 2% among SACE Physics teachers and the 31% of participants who moderated exceeded proportion of 17% of Physics teachers. Involvement of participants in the quality management processes reinforced the conclusion made earlier that the sample was not representative of SACE Physics teachers (see 4.1.1.4).

### **4.3 Reasons for teacher involvement in the quality management processes**

The reasons for teacher involvement, or non-involvement, in the quality management processes were elicited through teacher interviews, although officers also reported what teachers had told them.

#### **4.3.1 Reasons for involvement in planning forums**

Attendance at planning forums had been because participants wanted to learn about the changes in assessment, particularly the use of performance standards. Kevin explained: “It was very much in my interests and the interests of my students to go along and find out as much about it as I possibly could”. Another had expected to find out the reasons for the changes being introduced and three had expected forms and exemplars of the LAP and assessment tasks to be available. Harry attended to give feedback to the SACE Board, hoping to influence the changes that were being introduced. The most common reason for planning forum attendance was to learn about the changes in assessment.

#### **4.3.2 Reasons for involvement in LAP approval and clarifying support panels**

For one of the two supervisors, 2011 was a continuation of a previously-held role so there was no conscious decision to be involved in the new quality management processes. The other undertook the role after expressing reluctance. Although experienced in examinations, the person was conscious of lack of moderation experience. Realising that this applied to all teachers of the subject, the person agreed to be involved, reasoning: “I’ve got the [assessment] experience” but indicated to SACE Board officers the need for “quite a bit of training”.

Three participants did not apply for the LAP approval panel because they felt unqualified. Brenda had no previous SACE panel experience and was unwilling to be involved, citing lack of time. The other seven had SACE panel experience but did not volunteer for LAP approval panels. Colin, who alternated between IB and SACE involvement, was “in one of the IB years”. Two had school commitments at the time of the approval process and another two could see no benefit in approving LAPs. Three participants had decided to reduce SACE Board involvement due to frustrations in the preparation for the new SACE, believing that teacher contributions

had been ignored or devalued. Kevin said, “if they’re not going to listen then there’s no point me doing it. I’ve got better things to do with my time.” Frances reported: “I’ve got my back up a little bit about the whole process, ... how everything has been brought in without them listening to us.”

Participants involved in LAP approval were subject supervisors. Others declined involvement because of inexperience, other commitments or frustration with the new SACE implementation.

### **4.3.3 Reasons for attending clarifying forums**

Teacher attendance at clarifying forums might be based on the SACE Board’s descriptions of their purpose, described as providing opportunities for teachers to “receive feedback on the school’s application of the performance standards” (SACE Board of South Australia, 2010i, p. 5) and supporting “teacher discussion on assessment task design and assessment of student work against the performance standards” (SACE Board of South Australia, 2010m).

Participants had attended expecting to learn. Clarifying was “the key word” for Russell, who went in expectation of “getting a lot of questions answered”. Alina described their purpose as “adjusting the way we mark, so we get the right answer”. Adam, reflecting later, suggested “they were trying to get us to think about marking in a different way.” A number of participants anticipated improved understanding of the performance standards and clarification of the moderation process.

A number showed eagerness for information about their own practices, with Alina hoping “to confirm that what I’m doing is right”. Brenda, wanting to ensure that she was “on the right track”, took some of her tasks for feedback from other teachers. John had delayed assessment until the clarifying forum “just to make sure we’d got it right”.

Participants attended clarifying forums wanting to learn about the use of performance standards in task design, grading student work and to get feedback on their own assessment practices.

#### 4.3.4 Reasons for involvement in moderation

Moderators see moderation as an important part of their professional development (B. Atherton, 2009; Maxwell, 2006). It was, therefore, anticipated that Physics teachers might apply to moderate for the learning opportunities afforded, particularly in the first year. The information sheets on panel membership did not mention benefits but membership was described as an “opportunity” (SACE Board of South Australia, 2010n), generally carrying the connotation of being advantageous in some way. In information on moderation panels, mention was made of the training that moderators would receive (SACE Board of South Australia, 2010m).

Participants who applied described the benefits in terms of the learning that might occur. Harry, with SACE Board panel experience, described previously feeling “like I was a sparrow sitting on the wire, waiting to be shot off” and decided to “get involved.” Adam moderated “to see how it all works” while Peter, from previous moderation experience, said “it will be useful”. Frances described the benefit of finding out “what [moderators] value and what they don’t”. Alina expressed interest in attending because of the promised training and to see how decisions were made “on a full set of tasks”. Greg, self-described as “puzzled and confused”, was attempting “to understand the machinations of this mysterious process”.

Three did not apply to be moderators because of inexperience, while the others cited inability to be absent from school for five days. One participant with previous panel experience, who had mentioned frustration with the preparations for the new SACE, did moderate at the end of 2011 and others with panel experience were intending to moderate in 2012. Conversely, two who moderated in 2011 were rethinking their involvement because, exhausted after moderating all day, they had to mark school examinations at night.

SACE Board officers reported differences in teacher willingness to moderate. One officer explained how, in some subjects, panellists “do it because they have a passion for the subject, ... they want to learn and they also want to teach everybody else as well”. Others reported this was not always so; in other subjects, non-applicants were approached to find enough moderators. After moderation, officers suspected that finding sufficient moderators might subsequently be more difficult: moderators had “worked so hard ... that they really don’t believe that they want to be part of that

process again”. Officers suspected that some teachers, moderating for the first time “to find out what was going on”, were unlikely to repeat the experience. This was particularly so for teachers who had undertaken both examination marking and moderation and were “close to burnout”.

It appeared that all participants who volunteered for moderation did so for its anticipated educative value.

#### **4.3.5 Reasons for involvement in the quality management processes**

A few participated in the quality management processes so that they might provide teacher input into the evolving new assessment system. The most common reason for involvement, however, was the belief that they would increase their knowledge in using performance standards, learning described by participants as beneficial for their students. This is consistent with reports that teachers engage in professional development for the benefit of their students (Guskey, 1986). If teachers aspire to improving their knowledge and skills, and if involvement serves this end, it appears that strategies are required that maximise teacher involvement in the quality management processes.

Taking 2012 moderation intentions into account, by the end of 2012 eight participants would have experienced moderation. Of the other five, three believed they were ill-equipped and two had school commitments at that time.

Recommendations for increasing teacher involvement are discussed later (see 8.4).

#### **4.4 Teachers’ school characteristics and involvement in the quality management processes**

Statistical analysis of participants’ characteristics and involvement in SACE Board quality management processes was considered inappropriate because of the small sample. Data on teacher characteristics were available, however, from other sources.

Planning forum evaluation data (SACE Board of South Australia, 2010a, 2010b) reported that 25% of surveys were from Catholic schools, 55% from Government schools and 25% from Independent schools. These are similar to school proportions of Catholic 18%, Government 51% and Independent 31% (SACE Board of South Australia, 2011a). That 86% of survey respondents were from metropolitan schools,

while only 74% of SACE schools are metropolitan (SACE Board of South Australia, 2011a), implied that non-metropolitan teachers might have found attendance more difficult than metropolitan teachers. Officers were aware of non-metropolitan teachers unable to attend who had contacted them for the materials distributed at the forums.

Observation of LAP approval panels in three subjects suggested that sector distribution among panellists was fairly consistent, with approximately half from Government schools and with Catholic and Independent sectors each providing approximately one quarter, with one panellist from outside the Adelaide metropolitan area. It appeared that teacher involvement was determined by the SACE Board for “political” reasons, rather than teacher preferences, rendering analysis in terms of school location or sectors problematic.

Clarifying forum data (SACE Board of South Australia, 2011b) revealed similar high attendances from all three sectors and that the forums were also well-attended by non-metropolitan teachers.

#### **4.5 Professional and gender teacher characteristics and involvement in the quality management processes**

As in the previous section, the small sample size made it inappropriate to analyse participant involvement on the basis of professional characteristics.

Observation of three LAP approval panels suggested that all panellists were experienced teachers with prior experience on SACE Board assessment panels. Most panellists had at least fifteen years of teaching experience, a few approximately ten years. Where possible, female-male balance was evident, possibly because of SACE Board directives.

It was considered there were insufficient data to allow a valid conclusion regarding teacher characteristics and involvement in the quality management processes.

#### **4.6 Communities of practice and the quality management processes**

Membership of communities of practice was included in the current research because of its importance in teacher learning (Díaz-Maggioli, 2004; Lave & Wenger, 1991;



---

Wei et al., 2009) and the SACE Review (Crafter et al., 2006) hoped to encourage the development of such communities among teachers. Data were examined for evidence of teacher participation in the SACE Board Physics online forum, intra-school collaborations and inter-school collaborations.

#### 4.6.1 Participation in the Physics online forum

The SACE Board online forums provide teachers with opportunities to “ask questions, discuss ideas and share experiences about teaching practice in each SACE subject” (SACE Board of South Australia, 2013c). Since the invitation to participate in the current research was issued through the online forum, all participants were members.

As a member of the online forum, I received each notice. In October 2010, teachers began asking questions, making comments and responding to questions about the new SACE Physics course. Between then and the end of 2011, 57 notices were posted by 29 different teachers (approximately 20% of the SACE Physics teachers in South Australia) with eleven posted by SACE Board officers and one by the Science Teachers Association of South Australia (SASTA). One teacher posted eleven notices, one five, and another two posted four notices each; the other 25 teachers posted one or two notices.

Based on the purposes of the forums (SACE Board of South Australia, 2013c), postings were categorised into *ask question*, *discuss ideas* and *share experiences*. Another category *provide information* addressed other notices, including some where question responses were not conducive to discussion.

Of the 57 notices posted by teachers, 34 discussed ideas. Fifteen asked questions, two of which sought to combine small classes. Five provided information and teachers shared their experiences in the remaining three. Teachers appeared unwilling to share their experiences and were more inclined to ask questions and discuss ideas, which might be anticipated in the first year of new assessment practices.

In late March a teacher suggested a teacher-access-only site where teachers could post tasks and download others. Within twelve hours a SACE Board notice informed

teachers that the online forum was restricted to teachers: the forum had been established “to enable teachers to discuss and share materials” and files could be attached to notices. The notices anticipated “more active use of the forum”. Within two days one teacher indicated that an alternative site existed and another reported a site being established. SASTA also offered its website but indicated that access would be restricted to members. Two months later, the next notice on this thread was from a first-year teacher awaiting “any materials anyone has to offer”. A prompt reply suggested four reasons why teachers were unwilling to submit materials: they were unsure of “getting it right” in “uncharted territory”, they were unwilling to expose their work to criticism; they were wary of revealing that most tasks were unoriginal, and teachers feared that circulating their tasks would restrict their reuse.

All notices posted by SACE Board officers and SASTA provided information. The SACE Board did not use the online forum to stimulate teacher discussion. It might be argued, however, that it is the teachers’, not the SACE Board’s, responsibility to stimulate discussion and the Australian National Professional Standards for Teachers (Australian Institute for Teaching and School Leadership, 2011b) expect highly accomplished teachers to “Initiate ... professional discussions with colleagues” (p. 18). The SACE Board did use the online facilities in association with the clarifying forums. After enrolling, teachers were granted online access to samples of student work and invited to grade them and register their grades and comments prior to the forums. The grades and comments were not accessible to other teachers. All participants tried to download the work, one failing because of computer problems. Nine participants graded the work before the forums and three were among the 28 who registered their grades.

In 2012 fewer than 20 notices were posted in the Physics online forum, well below the 57 of 2011. It appeared that 2011 saw increased use of the online forum because of the changes but that the potential for discussion and sharing of materials had not been utilised.

#### **4.6.2 Intra-school collaboration**

Four participants were the sole Physics teachers in their schools and, while they occasionally met socially with other Physics teachers, they essentially tackled 2011 alone. One expressed frustration at a SACE Board suggestion to “Go talk to other

Physics teachers”, explaining: “It’s hard enough to talk to people in the school, let alone to find another Physics teacher.”

Eight participants (62%) were in schools with more than one Physics teacher. This meant participants had greater opportunities for intra-school collaboration than Physics teachers in general where only 34% are in schools with other Physics teachers (SACE Board of South Australia, 2011a). In all cases, the teachers collaborated throughout the year, with most continuing the practice of previous years. For one, previous collaboration was impossible because 2011 was her first year in the school. Thus, for the participants, intra-school collaboration was unaffected by the changes of 2011.

#### **4.6.3 Inter-school collaboration**

Inter-school collaboration intensified during 2011. The Association of Independent Schools of South Australia (AISSA) provided funding to allow teachers to form localised communities of practice, commonly referred to as *hub groups*. SACE Board officers reported “hubbing” in metropolitan and country areas in all three sectors and had worked with some groups, often by email response to questions raised by the group, although other groups had declined officer support. One participant recalled an officer attending an inter-school meeting.

Six participants were members of inter-school collaborations, four being members of both intra- and inter-school communities. All inter-school communities involved schools from the same sector and, in most cases, close proximity.

The data indicated that most Physics teachers tackled 2011 alone. In schools with more than one Physics teacher, teachers worked collaboratively, while the ad hoc inter-school collaborations were short-lived, existing specifically to assist teachers address the changes of 2011.

#### **4.6.4 Encouragement of communities of practice in the quality management processes**

The current research scrutinised the data for evidence that the quality management processes encouraged or facilitated communities of practice among subject teachers.

In planning forums and clarifying forums teachers sat in groups, more conducive to

discussion than rows of seats facing the front. One officer identified the opportunity for networking at the clarifying forums, while another recalled teachers exchanging contact information. A third detailed the generous responses of teachers to others' insecurities, inevitably the consequence of being working alone in their schools. Officers described teachers at clarifying forums agreeing to share resources. Three participants used the clarifying forums to get others' opinions on their assessment practices. While planning and clarifying forums created the possibility of teacher networks, it would be generous to claim that networking was encouraged.

The opportunities for teacher communication in the online forum saw increased use in 2011. Although teachers had advocated task sharing and the SACE Board had indicated the forum might be so used, this did not eventuate. It appeared that, although the SACE Board provided opportunities that might generate communities of practice, little was done to encourage or facilitate such engagement and intra- and inter-school collaborations arose from teacher initiatives, with some officer support.

#### **4.7 Use of exemplars**

Exemplars are one strategy used in quality assurance of SBA (Harlen, 1994; Hipkins & Hodgen, 2011; Sadler, 1987) and the SACE Board, while not including exemplars in their quality management processes, acknowledged a responsibility for their provision (SACE Board of South Australia, 2012f).

Twelve participants used the website exemplars to inform themselves about performance standards and the revised assessment procedures. Peter explained they were "useful if you were not quite sure about what the performance standards were talking about". Kevin acknowledged: "They gave us ideas about how to arrange assessments" and Harry changed his "prac write up to suit the exemplar". Others used the tasks as a guide for writing their own assessment tasks or discussed them with students.

Participants accessed the SACE Board website exemplars and found them useful in various ways, suggesting they might be used more widely, as examined later.

## 4.8 Overview of teacher involvement

Participant involvement in the quality management processes arose from the expectation of learning, although a few saw an opportunity to provide “a teacher’s point of view” on the change process. The changes of 2011 stimulated increased collaboration through the online forum and inter-school hub groups. Hub groups were, however, short-lived and there was a failure to maximise the online forum’s potential.

While there was substantial involvement of teachers in SACE Board activities during 2011, most of this was considered induction to the changes, rather than as part of an ongoing quality management cycle. Increased involvement in SACE Board assessment roles appeared minimal and did not appear to satisfy the call for greater teacher moderation involvement issued by the Review Panel (Crafter et al., 2006). Without substantial involvement, the goal of continuous improvement in teaching and learning appears unlikely.

A number of questions arise. Why did the new SACE not adopt the Review Panel’s recommendation for greater teacher involvement in social moderation? Should the SACE form of social moderation be changed to that recommended by the Panel? Does social moderation serve other purposes than teacher learning that justify its quality control role in school-based assessment, or is it dispensable? If teacher involvement in social moderation is restricted, are there other means of delivering its educative benefits to non-moderators? These questions are explored later.

The next chapter analyses the form of the SACE Board processes in the light of the literature on effective teacher learning, exploring further the potential for teacher learning in quality management processes.

## **CHAPTER 5: RESULTS: FORM OF THE QUALITY MANAGEMENT PROCESSES**

In this chapter the SACE Board’s quality management processes are examined for evidence they were consistent with those identified in the literature on effective professional learning. In particular, the quality management processes were examined in the light of the literature on professional development activities, teacher learning, adult learning and the encouragement of deep learning.

### **5.1 The quality management processes as effective professional development activities**

In this thesis, professional development activities refer to the one-off activities specifically designed to enhance teacher professional knowledge and skills. For the SACE in 2011, these included the planning and clarifying forums and the assessment panel training sessions: learning and assessment plan (LAP) approval, clarifying support and moderation (SACE Board of South Australia, 2010m). Table 23 summarises the inferences when these were examined for the features of effective professional development activities: active learning, coherence, duration and collective participation.

**Table 23: Overview of quality management processes as effective professional development activities**

<b>Process</b>	<b>Active learning</b>	<b>Coherence</b>	<b>Duration</b>	<b>Collective participation</b>
Planning forums	Yes	Partial	No	Yes
LAP approval panels	Yes	Partial	No	Yes
Clarifying support panel	Yes	Yes	Yes	Yes
Clarifying forums	Partial	Partial	No	Yes
Moderation panels	Yes	Partial	Yes	Yes

The following pages describe the evidence supporting these conclusions.

### 5.1.1 Active learning in the quality management processes

In the context of assessment, construction and analysis of assessment tasks and the review of student work were considered to be active learning. The quality management processes were examined for evidence of such activities.

The planning forums introduced teachers to the changes of 2011, with a focus on using performance standards in assessing student learning. Grading of student work was undertaken with performance standards, which was considered to constitute active learning since it involved meaningful analysis of an element of teaching and learning (Garet et al., 1999).

Four LAP approval panel sessions were observed in three different subjects. After an introduction by the relevant SACE Board officer, panellists were led through one approved LAP. Discussion was followed by consideration of another LAP for suitability, followed by further discussion. The session continued with panellists determining the approvability of schools' LAPs. It was concluded that active learning was involved in the process.

Clarifying support panels consisted of the three subject supervisors. Observation of two panels (in different subjects) identified opportunities for active learning in the discussion of task design and the grades to be awarded for student tasks. Two participants served on clarifying support panels. One participant described "working with the performance standards and firming up in our mind what the different things were" and becoming aware of flaws in the task design. It appeared that the discussions, involving meaningful analysis of assessment tasks and the application of the performance standards to student work, could also be considered active learning.

For the clarifying forums, samples of student work were put online. Teachers were invited to download and grade the work and register the grades online, prior to the forums. Twelve participants downloaded the work, nine graded it and three registered their grades. Observation and interviews confirmed that discussion of the grades assigned to student work was undertaken. The value of the discussion appeared questionable, however. There was agreement among presenters and teachers that the design of the assessment task was flawed. Participants, however, expressed frustration at the lack of explanation. As one said: "You didn't get any

---

guidance on how to make things better.” In general, participants recalled unhelpful discussion on the reasons for the grades assigned. Brenda described it as “a very negative experience ... [because] I don’t think it was explained very well” and Russell “never found out what I was doing wrong”. It is questionable whether the clarifying forum discussion could qualify as “meaningful analysis” (Garet et al., 1999) since participants were uncertain about the rationale for grading decisions.

The six participants who were moderators reported active learning in the moderation panel training. Panellists were provided with a sample of student work and asked to grade it. The grade previously assigned by the supervisors was then revealed and discussion followed. This process was repeated with two other student tasks, except that panellists were told the assigned grade when each sample was distributed.

Participants reported frustration: “It never seemed to get resolved ... why it was a C.” A number would have preferred the discussion to have been “a lot longer to get us to be more precise with our judgements”. After the grading training, panellists were instructed in administrative procedures: identification of student work for scrutiny in each class, completion of the results sheets and the writing of feedback. In these processes panellists practised with class lists and results sheets. The moderation was conducted by panellists working in pairs which changed daily. Moderation involved scrutiny of selected student work and discussion within the pairs. Where grade changes were considered necessary, supervisors were consulted. It was considered that active learning was involved in moderation because it involved dialogue with peers (Zemke and Zemke, 1988) and “meaningful analysis” (Garet et al., 1999, p. 3-5) of assessment tasks and student responses. Later discussion includes providing access for more teachers to the effective teacher learning associated with involvement in social moderation.

Although effectiveness varied, the evidence suggested that all five one-off quality management processes provided teachers with experiences of active learning and, through enhanced discussion opportunities, had the potential for greater learning for teachers. Participants’ reports of these learning opportunities are addressed later (see 7.1).



### 5.1.2 Coherence in the quality management processes

Coherence of professional development activities involves connection with previous and subsequent activities, alignment with relevant subject documents and the encouragement of ongoing professional communication (see 2.6.2.2). The quality management processes were scrutinised for evidence of coherence: content that was common to different quality management processes.

The planning forums, conducted to support teachers in making changes to their assessment practices (SACE Board of South Australia, 2010d), were linked to previous activities through reference to the syllabus. Alignment of assessment tasks with the syllabus and the links between learning requirements, assessment design criteria and performance standards were apparent. Grading of student work preceded the requirement that teachers undertake such grading in 2011. The forums were conducted with teachers of the same subject arranged in groups around tables, which facilitated discussion among peers. There was no evidence, however, of encouragement of professional communication beyond the forum session. It seemed, therefore, that the planning forums demonstrated coherence in many ways, but did not foster ongoing professional discussion.

LAP approval appeared to satisfy the requirement of coherence because it linked to preceding activities (e.g., planning forums) and subsequent activities (e.g., moderation) and to relevant SACE Board documents (e.g., syllabus). Ongoing professional communication with others, however, was not evident since panellists mostly worked alone and there was no “ongoing” discussion after the approval completion. It seemed, therefore, that the LAP approval panels demonstrated only partial coherence.

The clarifying support panel involved selection of student work followed by intensive discussion on task design and the grades to be assigned. Since the panellists had worked together on LAP approval and did so in moderation, there was ongoing professional communication. Consequently, it was considered that the clarifying support panel demonstrated coherence.

In clarifying forums, teachers were presented with materials prepared by the support panel. They had the opportunity to grade the student work and discuss the materials.

---

Again, it was considered that by failing to foster ongoing professional discussion, the clarifying forums demonstrated partial coherence.

Moderation demonstrated coherence with previous activities since moderators were subject teachers who had undertaken student assessment during the year and submitted student work for moderation. Moderation was not, however, envisioned as part of an ongoing process; moderators were appointed for the five-day moderation process. Hence moderation was considered to demonstrate partial coherence.

The quality management processes strongly demonstrated two aspects of coherence: connection with previous and subsequent activities and alignment with relevant subject documents. Except for the supervisors, however, the processes did not encourage ongoing professional communication. The SACE Board did not attempt to continue professional discussion among teachers beyond the occasions when teachers met—in planning and clarifying forums and moderation meetings. While they made available the online forum, they did not use this to stimulate discussion among teachers. Modifying the quality management processes to facilitate continuing teacher dialogue would enhance their coherence and, consequently, their potential for effective teacher learning.

### **5.1.3 Duration of the quality management processes**

For the current research, 25 contact hours and a time span of 4 days was established as necessary for effective professional development activities (see 2.6.2.1). The quality management processes were examined for compliance with these specifications.

The analysis of teacher involvement (see 4.2.5) identified three levels of involvement among Physics teachers: non-moderators, moderators, and supervisors. For those who were neither moderators nor supervisors, involvement was restricted to attending planning and clarifying forums: approximately six contact hours over two days. These teachers did not have the minimal necessary contact (i.e., 25 contact hours and a time span of 4 days) described above as necessary for effective professional development activities.

Moderation involved approximately 35 hours over four days. Most moderators also

---

attended planning and clarifying forums, so their total contact time exceeded 40 hours over six days, satisfying the duration specified.

Supervisor involvement was substantial. In addition to moderators' commitments, they spent approximately six hours on LAP approval, six hours preparing for clarifying forums and six hours leading clarifying forums. Thus commitment in 2011 exceeded 50 hours over twelve days, throughout the year.

Although moderators and supervisors experienced the necessary duration for effective professional development, the majority of Physics teachers did not have this opportunity. Extending the quality management processes to provide opportunities for greater duration of teacher engagement would enhance their potential for effective teacher learning. Recommendations for increasing teacher engagement in the quality management processes are discussed later (see 8.4).

#### **5.1.4 Collective participation in the quality management processes**

Participation of groups of teachers from the same school, subject or grade has been found to contribute to the effectiveness of professional development activities (Garet et al., 1999; Garet et al., 2001).

In the metropolitan area, planning forums were generally held for teachers of one subject. In country areas and for a few city meetings, forums involved teachers of similar subjects (e.g., Chemistry and Physics); different materials were used for each subject and discussions were among teachers of the same subject.

The LAP approval panellists were teachers of the same subject and clarifying support panels consisted of three teachers of the same subject. The clarifying forums generally involved teachers of one subject although, in some country areas, forums involved teachers of similar subjects (e.g., Chemistry and Physics). Moderation panels consisted exclusively of teachers of the subject being moderated.

The quality management processes exhibited one form of collective participation; they involved teachers of the same subject and year level. This was one of a number of features of the quality management processes that showed their potential for teacher learning. Recommendations for making more effective use of this potential are discussed in Chapter 8.

---

### **5.1.5 Conclusions about quality management processes as effective professional development activities**

Overall, the one-off quality management processes were consistent with many aspects of effective professional development activities: active learning, coherence and collective participation. They did not, however, encourage ongoing communication between teachers.

For most teachers, the chief limitation of their effectiveness lay in their duration. For three supervisors and 25 moderators, the experiences of 2011 appeared consistent with the necessary duration of effective professional development activities. For the majority, however, the restricted time spent on the processes limited their effectiveness.

The data indicated that the quality management processes have the potential for effective teacher learning. This potential might be better realised by modifying them to encourage ongoing communication among teachers, thus increasing teacher engagement (see 8.4 for recommendations in this area).

## **5.2 Features of effective teacher learning in the quality management processes**

Recent literature has focussed on other forms of teacher learning than the one-off activities considered above. (see 2.6.3.2). The quality management processes were examined for evidence that they afforded learning opportunities, encouraged intra- and inter-school moderation during the year and provided access to expertise.

Evidence was also sought concerning the extent to which teachers' opinions in the area of assessment were challenged. The following sections scrutinise the quality management processes for evidence that they might have possessed these features.

### **5.2.1 Learning opportunities embedded in teachers' work**

The quality management processes engaged teachers in approving LAPs, discussing assessment tasks and grading student work. These activities, although similar to the teachers' work, were not embedded in it; the LAPs, assessment tasks and student work discussed were not their own. The activities might have afforded insights into their own work but the teachers were expected to make connections between the activities and their own work.

---

Participants reported different levels of discussion concerning their own work in their schools. Sole Physics teachers in schools were restricted to implementing understandings gained from the quality management processes. Others reported ongoing discussions with Physics teachers in their schools. They were able to discuss their understandings and apply them to their own situations. Suggestions were tossed about, tasks were trialled and reviewed on a regular basis. Such intra-school activity would appear to comply with the concept of teacher learning embedded within practice.

There was no evidence that the quality management processes encouraged practice-embedded learning in the area of assessment. Where practice-embedded learning occurred, it was because teachers in the same school worked collaboratively. The potential for teacher learning in the quality management processes might be improved were they to provide similar opportunities for the sole subject teachers in their schools.

### **5.2.2 Moderation opportunities during the year**

Teachers need time to develop common understandings of performance standards (Hipkins & Robertson, 2011; Klenowski & Wyatt-Smith, 2010b; Smaill, 2013; Wyatt-Smith et al., 2010), with little opportunity for such learning if moderation is limited to one annual event. Additional opportunities might be available to teachers through intra- or inter-school moderation during the year, ongoing grading discussions among teachers or by providing feedback to teachers on tasks and grading during the year. The quality management processes were examined for evidence.

SACE documents were scrutinised for evidence they encouraged intra- or inter-school moderation during the year. A consultation paper (*future* SACE Office, 2008) mentioned both intra- and inter-school moderation, although intra-school moderation would only be necessary if classes undertook different assessments, by tackling different tasks or having work marked by different teachers. The 2010 moderation procedures (SACE Board of South Australia, 2010i) required schools to join to “ensure ... a minimum of ten students” (p. 5) in each moderation group. Where this occurred, schools were expected to “conduct confirmation activities [i.e., moderation] ... to ensure that assessment decisions ... are reliable and fair across

---

classes” (p. 5). In 2011 schools were “strongly advised” in some subjects and “encouraged” in others (SACE Board of South Australia, 2011f, p. 101) to form assessment groups of ten or more students. Teachers were advised that they “*may* [italics added] share a common learning and assessment plan” (p. 102) or use a different plan for each class. Advice was given on procedures for ensuring “a common interpretation and application of the performance standards” (p. 102). Rather than encouraging intra- and inter-school moderation, it appeared that the eventual procedures allowed schools to avoid them if they wished—they were not mandatory.

Although discussion might contribute to common understandings, there is a risk of diversity rather than unity if discussion occurs in isolated groups. Dialogue is necessary between groups for common understanding to emerge. SACE Board officers commented on differing interpretations between different hub groups of teachers of the same subject. One commented that some became “quite aggressive” and the officer felt obliged to arbitrate. Another observed that a hub group representing a large number of schools had the potential “to dictate to the rest of the [subject] community”, with some hub groups “really closed” to new members. One officer saw isolated groups as exposing themselves to risk without “some sort of cross-pollination” or contact with SACE Board panels or officers. Some hub groups contacted officers for clarification and participants expressed satisfaction with officer willingness to respond.

It would appear that the SACE Board might be in a position to facilitate ongoing discussion among teachers, contributing to the development of shared understandings. There was no evidence of the Physics online forum being used to stimulate discussion among teachers. Previously, SSABSA had employed teachers and Board officers as *support moderators* (SSABSA, 2006) to provide advice to teachers prior to moderation. Their demise in the new SACE had been predicted by subject teachers (B. Atherton, 2009) and they were not included in the new system (*future SACE Office*, 2008). SACE Board officers commented that, in subjects with a history of support moderators, they had been “really missed”. Although Physics had no support moderators, three participants commented on their assistance in other subjects.

---

One task undertaken by support moderators was providing pre-moderation feedback on task design and the grading of student work. Frances commented that, since task design could have a profound impact on moderation outcomes, “the tasks should be checked before the kids sit them”. Completion of LAPs required minimal information on task design: feedback addressed specifications with no comments on task design.

Participants created opportunities to obtain others’ views on their assessment tasks at clarifying forums. Greg had contacted another Physics teacher for feedback on his assessment tasks, possibly because the two knew each other from SACE Board panel experience. Similar reviews for many teachers would require provision of time and some means of facilitating communication between teachers. SACE Board officers reported that teachers asking to submit tasks for comment was “quite common”. The officers’ responses depended on the time they had available when each request was made. One officer believed that it was not the Board’s role to respond to teachers “at the individual task level”; it was “inequitable to do it for some teachers and not for others”. Others worked with teachers, providing that the task be made available to others as support materials. Sometimes officers passed requests on to members of the relevant Curriculum Leaders Group.<sup>33</sup>

It appeared that, in the new assessment system, the opportunity for moderation during the year was reduced. Some teachers initiated review of their materials prior to moderation and SACE Board officers were willing to respond to approaches made at convenient times. Providing moderation opportunities during the year, for example through support moderators, and improving teacher learning potential of the quality management processes, is examined later.

### **5.2.3 Access to expertise**

Support for teachers in times of change is important (e.g., Fullan, 2011; Hipkins & Robertson, 2011; Timperley et al., 2007). The degree of support for the assessment changes described by participants varied widely. This section examines the availability of experts, based on observations, focus groups and teacher interviews.

---

<sup>33</sup> Curriculum Leaders Groups exist in most subjects to provide advice and contribute to teachers’ professional learning (SACE Board of South Australia, 2014d).

The leaders of the forums and assessment panels were SACE Board officers. Each officer who took part in the focus groups was a teacher with assessment experience in at least one Year 12 subject. Some had experience in a second and, rarely, a third subject. The forums were developed by officers, who presented forums in their teaching subjects and in others in which they did not have teaching expertise.

Planning forum exit surveys (SACE Board of South Australia, 2010a, 2010b) did not specifically ask about presenter expertise. Although data were not available for Physics teachers, science teachers' responses showed satisfaction with the presenters: 72% reported the subject-specific content as good or very good and 74% gave similar ratings for the overall presentation of the forum. Thus, it appeared that teachers were satisfied with presenter expertise.

Participants, aware that SACE Board officers advised in subjects in which they had no teaching experience, responded differently. One participant objected to training being provided by a non-Physics presenter, but others recognised that the SACE Board could not employ experts in all subjects. Kevin, knowing officers were working outside their teaching subjects, acknowledged their assessment expertise, referring to them as "SACE experts". The majority expressed satisfaction with the expertise of the presenters, although one commented on their inexperience with performance standards, common among SACE teachers as well. The supervisor-participants who led clarifying forums described themselves as "trying to stay a step or two ahead" and "making it up as we go". Although classified as experts, they did not appear comfortable in the role.

Two participants spoke of a number of emails and phone calls exchanged with SACE Board officers. Colin described the interchange as useful, although some answers were described as "couched in political speak" which required "reading between the lines". Russell described how officers "get back to you as quickly as they can", adding: "When you do get onto them they're good." The SACE Board officers felt that limited time restricted their availability to work with teachers.

It appeared that participants acknowledged the officers as experts, even outside their teaching experiences, and found them to be supportive, subject to availability.



## 5.2.4 Challenging prevailing discourses

Effective teacher learning requires teachers' beliefs to be challenged when they vary considerably from what is being promoted in the learning (Thompson & Zeuli, 1999; Timperley et al., 2007): dissonance needs to be created and opportunities provided for its resolution. The data were examined for evidence of dissonance between teachers' beliefs and the practices they were required to implement. The quality management processes were scrutinised for evidence that they challenged teachers and provided them with opportunities to resolve any dissonance.

### 5.2.4.1 *Evidence of dissonance between teachers' beliefs and the new assessment practices*

The first round of teacher interviews revealed that all participants preferred the previous system, which they believed was working satisfactorily in Physics. They were required to implement a new system with insufficient knowledge and little faith. There was a dissonance between the previous system, which participants believed worked, and the new practices they were required to implement.

At the beginning of the research, seven commented favourably on aspects of the new system. One suggested the new system might be fairer: "at least student tasks are being looked at". Two acknowledged that the previous system had allowed schools to avoid practical work and two questioned the validity of using the examination to statistically moderate dissimilar school-assessed work, such as practical tasks. Two acknowledged the performance standards were useful: John said they forced him "to think about things that were 'on the edge'", while Peter stated they "made me think more about what an A standard is like".

Six participants could see no benefits in the changes. Kevin described them as "change for the sake of change". The additional time required by teachers to use performance standards was mentioned by six participants, with one describing assessment as "interfering with the learning"; more time was spent "ticking boxes" than teaching. Difficulty in reaching agreement on interpretation of performance standards was mentioned by six participants, consistent with the need for discussion among teachers (Adie, 2008; Crafter et al., 2006; Daugherty, 1997; QSA, 2009). John identified a limitation of performance standards—their inability to "encapsulate all about student performance"—similar to the observation that focussing on the

---

constituent parts can cause loss of sight of the whole (Harlen, 2007). No other participants made similar references to limitations of performance standards.

Six participants expressed greater faith in the “objective” nature of the previous system. One reported that, at the planning forum, teachers had been able to agree on the grade assigned to student work without recourse to performance standards; it was only when the standards were introduced that differences arose. Frances described their use as a “hand-wavy way of working out what an A+ might be”, while Russell claimed, tongue-in-cheek, “we should be able to hum and see what that averages out”. Six questioned the fairness of the new system on three grounds: comparability of assessment conditions in different schools, excessive parent or tutor support and the questionable reliability of social moderation.

In the current research, it was considered that some aspects of dissonance could not be addressed by quality management processes, for example the time required by teachers. On the other hand, the processes might provide opportunities for resolution of other aspects of dissonance. Thus the research examined the quality management processes for evidence that they might address such matters as the benefits and perceived shortcomings of performance standards and social moderation, as well as the difficulties of verifying student work.

#### **5.2.4.2 Addressing the dissonance**

It was considered that the only quality management processes activities that afforded opportunities for challenging teachers’ assessment beliefs were the planning and clarifying forums and the clarifying support panels. Other processes did not provide opportunities for question and discussion—considered necessary strategies for resolving dissonance (Loucks-Horsley et al., 2003).

In reporting planning forums, ten participants could not recall any explanation for the changes and John thought presenters were “reading the party line [and he] went away not necessarily convinced that they believed it”. Another described the changes as “foisted” on teachers who did not see a need for change. Five believed that the changes were politically motivated to improve SACE completion rates and that the focus was on the weaker students: a focus on “participation and not rigour”. SACE Board officers were aware of the discord, with one reporting “there were still people

---

walking out the door having arguments”.

Observation of planning forums identified opportunities for discussion and participants recalled “plenty of opportunities for questions”. It did not appear, however, that these discussions had diminished any dissonance. Comments such as: “Well there’s not much point because it’s not going to change anyway”, suggested that participants might have been seeking resolution by the changes disappearing, rather than appreciating their adoption. Russell’s recollection that “they were just trying to give the teachers a reassurance that they would cope” suggested that, rather than resolving dissonance by addressing it, the planning forums tried to ignore it.

SACE Board officers’ comments support the suggestion that challenging existing beliefs was avoided. They described the role of the forums as reassuring teachers that “it’s not going to be a scary thing”. This appears to contradict arguments that standards-based reforms “press for tremendous changes in classroom instruction” (Spillane, Reiser, & Reimer, 2002, p. 387). Although teachers contrasted the “objective” nature of the previous assessment with the “subjective” nature of performance standards, this was ignored in the forums. Officers described teachers arriving at clarifying forums nine months later “still having no clue”, suggesting the planning forums had placated teachers, rather than challenging them and providing opportunities to work through any dissonance.

Discussion among clarifying support panellists provided opportunities to clarify the changes being introduced with the relevant SACE Board officer. Both supervisor-participants described how their views had changed through such discussion as they became “more in tune” with the performance standards. Membership of clarifying support panels appeared to provide opportunities to address any dissonance.

Participants were generally dissatisfied with the clarifying forums. While questions were raised, John was “not necessarily sure that they were answered”. Inconsistency in the grading of an exemplar between presenters and attendees generated frustration: “And nothing was done to resolve that.” John described the defensiveness on the part of the SACE Board as “reasonable” because the Board was responsible for implementing decisions not made by them but “within the SACE Review and ... in Parliament”.

---

It would appear that, while opportunities were presented in the clarifying forums for questions and discussions, the outcome was not the resolution of dissonance but rather its aggravation. It appeared that only teachers serving as supervisors had the chance to learn by having their existing beliefs challenged and being given extensive opportunities for discussion, with implications for teacher acceptance of the changes.

### **5.2.5 Overview of the quality management processes as effective teacher learning**

The revised assessment practices required teachers to use performance standards in designing assessment tasks and in grading student work. Hindrances to the development of teachers' assessment beliefs and practices lay in the failure to challenge existing assessment beliefs and limited opportunities for feedback during the year. Failure of the induction program to challenge the existing assessment beliefs of teachers has implications for teacher implementation of change (Hargreaves et al., 2001; Timperley et al., 2007). Access to feedback was available to those courageous enough to ask but it depended on officers having time available.

It was considered that the quality management processes only partly demonstrated the features for effective teacher learning. Their focus on student work offers much potential for teacher learning (Little, 1999; Sykes, 1999; Timperley et al., 2007), potential which might be better realised if the processes were modified with teacher learning as a goal (see 8.4). In this way attention is likely to be paid to addressing the hindrances to teacher development by confronting conflicting beliefs and providing teachers with feedback on their assessment practices—using performance standards in designing assessment tasks and grading student work—during the year.

## **5.3 Features of effective adult learning in the quality management processes**

The literature review identified four criteria for effective adult learning: consultation in the planning process, relevance to learners' needs, congruence with prior experience and the use of appropriate activities and materials (Knowles et al., 2005; Loucks-Horsley et al., 2003; Zemke & Zemke, 1988). Table 24 shows the conclusions when the quality management processes were examined against these criteria.

**Table 24: The quality management processes as effective adult learning**

<b>Process</b>	<b>Input into planning</b>	<b>Relevance</b>	<b>Congruence</b>	<b>Appropriate materials and activities</b>
Planning forums	No	Yes	Yes	Partial
LAP submission and feedback	No	Yes	Yes	Partial
LAP approval panels	No	Yes	No	Partial
Clarifying support panel	No	Yes	No	No
Clarifying forums	Yes	Yes	No	No
Moderation panels	Yes	Yes	Yes	Yes
Moderation feedback	No	Yes	Yes	No

The following pages describe the evidence that led to these conclusions.

### **5.3.1 Quality management processes: teacher input into planning**

The quality management activities were planned by SACE Board officers, using teacher feedback from the planning forums.

Officers commented that, at the planning forums, they “weren’t allowed to say a lot of things”. They observed that the common format meant that teachers of some subjects were uninterested in some topics: statistical information “was just a waste of half an hour”.

Supervisors were the only teachers involved in planning; they selected and graded the student work used in the clarifying forums and for moderation training.

Consequently, it was considered that teacher input in the planning of the quality management processes had been minimal.

### **5.3.2 Quality management processes: relevance to teachers’ needs**

With the new assessment system, Physics teachers submitted LAPs for SACE Board approval and samples of student work for final moderation. They needed to use performance standards in designing tasks and in grading student work. Relevance of

the quality management processes depended on their informing teachers about performance standards and educating teachers in their use.

Planning forums supported teachers in developing LAPs and assessment tasks, and in using the performance standards to assess student work (SACE Board of South Australia, 2010d). All these were relevant to teachers. Observations and interviews confirmed that the performance standards were introduced, as was an overview of the new assessment procedures. Teachers discussed an example of a LAP and used the performance standards to grade student work. Attendances were lower than for clarifying forums and the timing of the planning forums appeared to affect their relevance. SACE Board officers mentioned teachers commenting that the planning forums had been too early, which may have restricted teacher recognition of their relevance. Although participant opinions varied on their timing, it appeared that the planning forums were relevant to teachers; they stimulated teachers to begin preparations for 2011 and introduced them to the changes.

Preparation and submission of a LAP was seen, on the whole, as relevant to participants' needs. Most disagreed that the LAP was merely bureaucratic, with only two arguing they were "filling it in to keep the SACE people happy." Their relevance lay in the opportunity to plan assessment tasks and select the most appropriate assessment design criteria and specific features for each task. One participant described their relevance as an assessment guide, particularly for inexperienced teachers but many participants with decades of teaching experience appreciated the benefit of documenting their practices and having it approved by the SACE Board. Colin and John reported that it made them think about what they were doing" and Harry said it helped teachers "to structure their assessment process". There was, however, some frustration with non-approval for "trivial things ... [that were] easy to administer". The data indicated that the participants saw the LAP approval process as relevant to their needs, although frustrated by aspects of the process.

LAP approval panellists were teachers, who were required to submit plans for their own schools. As one of the participants said: "It does make you think about exactly what you're doing." Consequently, membership of the panel was considered relevant to teachers' needs.

---

Clarifying support panel membership provided opportunities for discussion, described by participants as beneficial in developing understanding of the performance standards and working with them in task design and grading student work. Provision of learning opportunities relevant to their teaching responsibilities, made the clarifying support panels relevant to members.

Participants attended clarifying forums to learn about using performance standards in task design and grading student work. Teachers had used performance standards to grade online tasks before attending, and at the forums discussed the grades assigned, which stimulated consideration of design features of tasks that would allow students to demonstrate the highest levels of achievement. It was considered, therefore, that the clarifying forums demonstrated relevance to teachers.

Moderation panel training included grading of student work followed by discussion, to develop consistency in applying the performance standards. Although some participants would have liked more training, the development of consistency and the reasoning behind grade determinations were relevant to teachers' needs.

To provide feedback to schools, moderators completed forms showing final grades and the basis of any adjustments. They identified instances where task design had limited the opportunity for students to demonstrate their learning in relation to the performance standards. It appeared that the information collected for schools was relevant to teachers' needs.

Although there were minor criticisms, the quality management processes were relevant to teachers' needs.

### **5.3.3 Congruence of the quality management processes with teachers' prior experiences**

Effective adult learning requires congruence between what is being learned and the learners' backgrounds. The quality management processes were examined for the extent to which they built on teachers' previous experiences.

The planning forums introduced teachers to the new SACE requirements: LAP submission as well as the use of performance standards in both task design and grading student work. The SACE Board faced a dilemma in preparing for the

---

forums: demonstrating congruence with prior experience while describing the changes required. Kevin recalled thinking: “We were pretty right, because it only really tinkered with what was already in place.” Russell reported them as giving the teachers “confidence to adapt to ... a fairly major change” and described the process as “clever” because “there was very little change in syllabus content. It was just the assessment that was the radical change”. SACE Board officers described the aim of planning forums: “to reassure people it’s not going to be a scary thing” Since the forums focussed on similarities between the new syllabus and the old, it seems reasonable to consider that the planning forums demonstrated congruence with prior experience. There was an attempt to have teachers perceive the changes as evolutionary, rather than revolutionary, to make the changes palatable.

The requirement to submit a LAP for SACE Board approval was not greatly different, as previously teachers had provided assessment plans to students. Only the requirement for SACE Board approval was new, as was the format. Thus there was congruence with prior experience in LAP preparation.

The clarifying support panels were where supervisors realised the incongruence with prior experience. The two supervisor-participants described the need to lead teachers as “trying to stay a step or two ahead”. Because of awareness of their lack of performance standards expertise, one admitted the need for “quite a bit of training” when accepting the position. Consequently, congruence with prior experience was very limited for supervisors: they had assessment experience, but not with performance standards.

The clarifying forums were held at least three months after teachers had begun using performance standards in assessing students. Between the planning and clarifying forums approximately twelve months had passed. Given the long interval and the limited information available during that time, it is questionable whether teachers felt that the clarifying forums connected smoothly with prior experience. It was at the clarifying forums that participants became concerned at the magnitude of the change imposed upon them. The reassuring “mantra” of the planning forums, that “the standards haven’t changed”, was contested at the clarifying forums. Eight of the eleven participants with prior Year 12 Physics experience believed that expectations had changed; the performance standards made it easier to get a C grade, but more



---

difficult to get an A grade, than previously. Most participants expressed frustration at grades assigned by presenters to student work and did not understand the rationale behind the performance standards nor how their use generated the grades assigned. In the forums, teachers sought direction or advice in using performance standards in assessment task design and were frustrated when that was not forthcoming. Although teachers sought examples of questions that assessed particular specific features, this was not provided. Clarification of the wording of performance standards was not provided and discussion that might resolve differences in interpretation was limited. It was, therefore, considered that the clarifying forums failed to confirm teacher expertise and build on prior experience. Instead, the forums alerted teachers to the need for change and to their inexperience with performance standards. The induction program was limited in its duration, failing to facilitate teacher learning as they engaged in the learning associated with their implementation of the new system: experiential learning. Modifications to the induction program are discussed in Section 8.5.

Participants anticipated that moderation would provide a learning experience. They had attended planning and clarifying forums and had submitted LAPs. They had designed assessment tasks and graded student work, using performance standards and had submitted work for moderation. Even participants who had not applied to serve as moderators believed “I’m sure it would do my teaching the world of good.” Clearly, moderation provided a learning opportunity related to their earlier experiences.

Most of the quality management processes clearly demonstrated congruence with previous experiences. That others, particularly the clarifying forums, were less congruent is not surprising, given that the introduction of performance standards and social moderation were “fairly major changes” for Physics teachers.

#### **5.3.4 Use of appropriate materials and activities in the quality management processes**

The quality management processes were examined for evidence of activities and materials that addressed real problems and included active learning and discussion, considered necessary for effective adult learning (Knowles et al., 2005; Loucks-Horsley et al., 2003; Zemke & Zemke, 1988).

---

Planning forums included a variety of activities: information was provided orally and in handouts, and there were opportunities for discussion and for teachers to ask questions. Presenters indicated aspects of the 2011 syllabus similar to that of 2010, while also specifying changes. Teachers used the new performance standards to grade student work. The activities and materials appeared appropriate for adult learning. One participant described how the planning forum “would make people think ... so that teachers would start contemplating planning ahead”, while another said they initiated teachers “trying to understand the implications” of the changes.

In the planning forum exit surveys (SACE Board of South Australia, 2010a, 2010b) over 70% of teachers reported the materials as good or very good. Eight months later, current research participants were more critical: they had expected exemplars would be available for LAPs and assessment tasks. Participants were critical that the tasks used for grading practice had not been written using performance standards; John described them as a “retrofit[:] ... previously-used assessment items were re-jigged to fit the performance standards and they didn’t necessarily fit properly”. Participants had expected tasks that had been trialled. They also described the planning forum as being delivered “at a policy level ... rather than the micro level” that teachers were seeking. The expectation of more explicit information was unfavourably contrasted with interstate experiences where “all of the hiccups [were] ironed out” before implementation.

Two factors may have contributed to the different appreciations of the exit surveys and the participants. First, teachers’ opinions may have changed in the nine months that elapsed between the planning forums and the interviews. Second, the participants, with greater experience of SACE assessment processes, may have been more critical than other Physics teachers. It appeared that, while teachers in general were satisfied, the participants were less so. Nevertheless, it was considered that the planning forums engaged teachers through a range of appropriate activities.

By the time that LAPs were submitted, one exemplar was available on the SACE Board website. Most participants used this in writing their own LAP. As Elena said: “if I base it on that I can’t really go wrong”. When implementing the LAP, however, she found that “some of the things ... were not appropriate”, which was not realised “until we were actually doing it”. Kevin commented “we didn’t make decisions ...

---

working out what's best for a particular assessment item" and Alina submitted the exemplar, praying that appropriate tasks could be found. Participant recollections suggested that, while the exemplar provided was adequate, some adopted it without sufficient thought. It appeared that teachers used the exemplar as a safe path in a time of uncertainty.

In LAP approval a checklist was distributed and discussed for the approval process. Copies of the syllabus assessment requirements were distributed and two exemplars were discussed before teachers began the approval process. While panellists worked alone, there was encouragement of discussion and this was plentiful. It appeared that materials and activities were provided that engaged panellists.

Clarifying support panels prepared materials for clarifying forums (SACE Board of South Australia, 2010m). Panellists were advised that they would receive training in assessment design and the application of performance standards (SACE Board of South Australia, 2010m). Participants described the process as evolving during the year. Panellists began not knowing "what we were looking for, ... sort of making it up as we go". One believed that the SACE Board was trying to produce resources but was "not sure what the resources are that they need." Panellists selected clarifying forum tasks without performance standards experience. As one said: "It's hard to ... sit down and pick a standard at the beginning of the year, using last year's materials that weren't assessed against these performance standards". Consequently, tasks believed satisfactory, were found to be flawed for use with performance standards. But it was too late; the flawed tasks were used in the clarifying forums. Supervisor-participants recalled that they felt under-prepared for their roles in leading clarifying forums; the training they needed "didn't really come".

After the clarifying forums, attendees from a sample of subjects were asked to complete a survey (SACE Board of South Australia, 2010n). In Science, Chemistry teachers were surveyed and over 80% of them had found clarifying forums useful (SACE Board of South Australia, 2010n), similar to the proportions in other subjects. Based on participants' interviews, opinions on the Physics clarifying forum materials appeared atypical. Physics participants were highly critical of the flawed tasks, describing how the practical task "gave [the students] everything" and pointing out that "answering questions is not enough [to merit an A grade]". Elena explained that

“you need to let [the students] introduce the complexity”. Although the presenters explained why the example was flawed, the participants were unforgiving: “they still haven’t got exemplars of work that is 100% right”. Frustration was very evident. Harry, having downloaded the exemplar for the clarifying forums, changed his practicals to match it, later learning that the task was inappropriate. Although the flawed tasks used in the Physics clarifying forums apprised teachers of the need for appropriately designed tasks, it would appear that the materials were counterproductive: “An exemplar should benchmark the standard that is required.” It is possible to sympathise with the presenters. They had to select exemplar tasks and grade student work while still novices in the use of performance standards. The implementation timeline imposed on the SACE Board gave rise to the use of flawed exemplars, at least in Physics, that generated frustration in teachers and hindered their acceptance of the new assessment system. A period in which some schools trial a new system was mentioned by a few participants who commented on this practice in other states. It would appear that such a procedure might avoid the problems that might have been a consequence of too-swift implementation.

Various materials were used to in-service moderation panels. Panellists were asked to grade samples of student work and presenters explained the grading decisions made earlier. Discussion of the decisions occurred. Although some participants would have preferred more discussion and more examples, the materials used in moderation panel training appeared appropriate.

Post-moderation feedback addressed task design and teachers’ grading judgements. Both moderators and teachers, however, considered the feedback too limited, suggesting that it did not serve the needs of teachers.

Overall the SACE Board quality management processes attempted to engage teachers through discussion and activities with student assessment tasks, although they were not always appropriate. Leaders of the implementation had limited experience with performance standards and felt under-prepared. Awareness of interstate practice involving trialling of materials resulted in condemnation of the new SACE implementation. Perhaps, however, criticism should have been directed, not at the Board, but at those who determined the timeline imposed on the Board. One participant referred to a SACE Board manager comparing the implementation to

“trying to fly an aircraft while you’re putting the wings on”.

### **5.3.5 Conclusions about the quality management processes as effective adult learning**

The data showed that the quality management processes were very relevant to teachers but, given the nature of the changes, could not be expected to be congruent with prior experience. Their greatest deficiency appeared to be the failure to adequately involve teachers in the planning. Although officers attempted to engage teachers with appropriate activities, the implementation timeline appeared to thwart development of appropriate materials. Consequently, the quality management processes were considered partially consistent with the qualities of effective adult learning, which might be improved with changes discussed later (see 8.4).

## **5.4 Opportunities for deep learning in the quality management processes**

The term *deep learning* is used in this thesis for learning which involves *making meaning* of new knowledge, *working with meaning* and *questioning assumptions* underlying assessment beliefs (Moon, 2000). The quality management processes were scrutinised for evidence that they might encourage deep learning through active learning, building on prior knowledge, confronting and eradicating teacher misconceptions and providing opportunities for discussion with other teachers and with presenters.

Some of these features have been discussed previously as contributing to effective professional development activities, effective teacher learning and adult learning. Consequently, previous conclusions have been incorporated into Table 25, which summarises the conclusions concerning opportunities for deep learning in the quality management processes.

**Table 25: Opportunities for deep learning in the quality management processes**

<b>Process</b>	<b>Active learning</b>	<b>Building on prior knowledge</b>	<b>Confronting and eradicating teacher misconceptions</b>	<b>Discussion with other teachers</b>	<b>Discussion with presenters</b>
Planning forums	Yes	Yes	No	No	Yes
LAP submission and feedback	No	Yes	No	No	No
LAP approval panels	Yes	No	No	No	Yes
Clarifying support panel	Yes	No	Yes	Yes	Yes
Clarifying forums	Partial	No	No	Yes	Yes
Moderation panels	Yes	Yes	No	Yes	Yes
Moderation feedback	No	Yes	No	No	No

The evidence for these conclusions is addressed in the following sections.

#### **5.4.1 Active learning**

Except for LAP and moderation feedback, the *active learning* data from Table 23 has been incorporated into Table 25.

Feedback provided to schools consisted of written information, not considered to be active learning. Consequently, active learning was only evident in those processes that involved teachers working directly with SACE Board officers.

#### **5.4.2 Building on prior knowledge**

Deep learning is encouraged by building on what students know (Biggs et al., 2011) and is, in this thesis, considered the equivalent of the effective adult learning requirement for “congruence with prior experience” (Díaz-Maggioli, 2004, p. 138). Consequently, data in the Table 25 column *Building on prior knowledge* were transferred from the congruence column in Table 24. The quality management

---

processes varied in the extent to which they were able to build on prior experience. Similarities to previous teacher practice were emphasised in the planning forums and end-of-year moderation was clearly linked to moderators' assessment experiences during the year. Other quality management processes, such as the clarifying forums, which alerted teachers to the radical nature of the change, had limited opportunity to build on teacher prior knowledge.

### **5.4.3 Confronting teacher misconceptions**

Biggs and Tang (2011) advised that deep learning was favoured by teachers “confronting and eradicating students’ misconceptions” (p. 27). Effective teacher learning required that teachers’ beliefs need to be challenged where there were “considerable gaps initially between prevailing discourses and the approaches being promoted in the professional learning” (Timperley et al., 2007, p. 73). These were considered equivalent and so the conclusions on addressing dissonance, stated earlier (see 5.2.4.2), were used to complete the column *Confronting and eradicating teacher misconceptions* in Table 25.

It appeared that the SACE Board quality management processes, apart from the clarifying support panels, did little to address and resolve teachers’ concerns and misconceptions regarding assessment processes. The major changes in the assessment system for Physics concerned the use of performance standards and the introduction of social moderation. For teachers to accept such radical changes, it is necessary that their previous understandings be addressed and conflicts resolved. For the implementation that was the subject of the current research, the time for such confrontation may have passed as the induction program has concluded and most teachers have minimal involvement in the existing quality management processes. Challenging existing understandings should, however, be addressed with teachers new to the SACE when they attend meetings introducing them to the assessment system. The need to address such issues in other induction programs is discussed in Section 8.5.

### **5.4.4 Discussion with other teachers**

Five of the quality management processes have been described as including discussion opportunities for teachers: planning and clarifying forums, LAP approval

panels, clarifying forum support and moderation. It was observed that teachers used the opportunities available. Rather than equate teacher discussion with encouragement of deep learning, it was considered that some analysis of the nature of the discussions was warranted.

Because the planning forums sought to reassure teachers (see 5.2.4.2) rather than have them think about the principles behind the changed assessment, discussions were unlikely to contribute to deep learning. LAP approval panel discussions, focussing on assessment specifications of the syllabus, were likely to be at a surface level rather than encouraging deep learning. Use of a checklist for making approval decisions would also encourage a surface level approach. Discussion in clarifying support panels was considered to encourage deep learning: participants reported discussion that clarified understandings of performance standards and task design. Clarifying forum discussions were observed to be “heated” and this was supported by SACE Board officers and participants. It was considered that the discussions might stimulate subsequent thinking and discussion among teachers, so encouraging the possibility of deep learning.

Moderation panel discussions, like those of LAP approval panels, focussed on syllabus understandings. Whereas LAP discussions focused on task specifications, moderation discussions focused on interpretations of the performance standards’ grading descriptors. Moderation discussions were extended and between more teachers. Consequently, moderation discussions were more likely to encourage deep learning.

Overall, it was felt that encouragement of deep learning in the quality management processes was limited to clarifying support panels, clarifying forums and moderation panels.

#### **5.4.5 Discussion with presenters**

Opportunity for learners to participate in discussion with teachers is conducive to deep learning (Biggs, 1999; Ryan et al., 2004). In the quality management processes, this equated to opportunities for teachers (the learners) to engage in discussions with the presenters (teaching about assessment).



Five quality management processes provided such opportunities: planning and clarifying forums, LAP approval panels, clarifying support panels and moderation panels. In all of these processes it was observed that the officers were approachable during, and after, the sessions. Russell verified that teachers made use of this opportunity, describing how “there was a queue five deep” wanting to talk to the presenters after one of the clarifying forums.

#### **5.4.6 Overview of opportunities for deep learning in the quality management processes**

It is suggested that the quality management processes, as a whole, were not designed to result in deep learning. Instead they appeared designed to equip teachers with a knowledge of the new procedures that teachers were mandated to implement. Encouragement of deep learning was limited, for most teachers, to one occasion—the clarifying forums. Although other processes demonstrated some features conducive to deep learning, particularly discussion among teachers and with presenters, these involved few teachers (the assessment panels) or were of a technical nature unlikely to stimulate deep understanding (e.g., LAP approval panels). If change is to be understood, accepted, internalised and implemented, it is necessary that all teachers be given opportunities to engage in deep learning about the reasons for the changes, their benefits and potential problems. In this way, improvement in teacher assessment literacy would be facilitated (DeLuca & Klinger, 2010; Kilvert & Mercurio, 2007). Addressing these shortcomings, as discussed in Chapter 8, would enhance the educative value of the quality management processes.

### **5.5 An overview of the form of the quality management processes**

The data indicated that the quality management processes possessed many features promising effective teacher learning, particularly relevance, active learning and discussion opportunities. The forums were relevant to teachers and provided them with an introduction to the changes required, although the limited involvement of teachers in their planning may have reduced their effectiveness and their short duration restricted learning opportunities.

For many, the quality management processes were somewhat remote from their classroom practices, with ongoing learning opportunities restricted to those

courageous enough to approach SACE Board officers or those fortunate enough to have another Physics teacher in the school. Ongoing communication among teachers did not appear to be encouraged, with a risk that diverse interpretations might have arisen in isolated teacher groups. It appeared that the quality management processes had not challenged teachers' assessment beliefs or provided sufficient opportunities for resolving differences, thus limiting the likelihood of deep learning. Instead, they focussed on equipping teachers to fulfil the requirements of the new assessment system. The use of flawed exemplars, possibly because of the implementation timeline imposed on the SACE Board, generated teacher frustration and hindered their acceptance of the changes. Possibly the greatest frustration of participants appeared to be in using performance standards, with little advice offered on how they might be used in assessment task design and grading.

Processes introducing teachers to the new SACE and those that comprise part of the annual cycle might facilitate teacher learning by addressing identified weaknesses, as recommended in Chapter 8.

The next chapter examines the data for evidence regarding the content of the quality management processes, and whether the content addressed had the potential to contribute to teacher learning in the area of assessment.

## **CHAPTER 6: RESULTS: CONTENT OF THE QUALITY MANAGEMENT PROCESSES**

This chapter examines the content of the SACE Board's quality management processes for evidence they provided opportunities for teacher learning in the area of assessment. The processes were examined for evidence that they focussed on assessment in a particular subject, informed teachers of the reasons for the assessment, and whether their content might have contributed to teachers developing knowledge and skills relating to the use of performance standards in school-based assessment (SBA).

SACE Board documents, observations, teacher interviews and SACE Board officer focus groups were used to analyse the processes in relation to these matters.

### **6.1 Subject focus of the quality management processes**

A core feature of effective professional development activities is that they should have a focus on subject knowledge (Garet et al., 1999), in this thesis taken as being a focus on subject-specific assessment.

In metropolitan Adelaide, quality management processes requiring teacher attendance (e.g., planning forums) were generally held for teachers of one subject whereas, in country areas, they involved teachers of different, yet allied, subjects (e.g., Chemistry and Physics). Where meetings involved more than one subject, different materials were used with each group and discussions were between teachers of the same subject. Assessment panels (e.g., clarifying support and moderation) always involved teachers of one subject and used materials specific to that subject.

Thus, the processes demonstrated the advantage, identified in the literature, of having a subject-specific focus.

### **6.2 Awareness of the reasons for change**

This section describes the evidence that quality management processes might have informed teachers of the reasons for the changes to the SACE assessment system. Data used were focus groups of SACE Board officers and teacher interviews.

Officers felt that teachers did not appreciate the reasons for the change, although they did believe that appreciation improved over the research period.

Most participants could not recall planning forums including an explanation for the changes. One participant suggested: “I think they were given by leaders reading the party line. I went away not necessarily convinced that they believed it.” The six months between the planning forums and the interviews would have contributed to participant failure to recall being told the reasons for the changes. When asked what they believed were those reasons, they gave a variety of answers, which are discussed later (see 7.2.1).

Observations, reading of materials distributed and participant interviews showed that subsequent processes focussed on implementation and offered no rationale for the assessment changes.

Participant failure to recall being told the reasons for the changes and the diversity in their explanations, suggested that reasons for the changes were not explained effectively. The induction period having passed, ongoing quality management processes might be used to enhance teacher literacy in SBA, as discussed in section 8.4.

### **6.3 Performance standards: purposes, advantages and disadvantages**

The introduction of performance standards was a major feature of the SACE assessment changes of 2011. It was essential that teachers be made aware of the purposes of performance standards, as well as their advantages and disadvantages. This section examines the content of the quality management processes for evidence that they might have informed teachers about performance standards.

Each SACE syllabus (e.g., SACE Board of South Australia, 2011d) states that the performance standards describe “the knowledge, skills, and understanding [used] ... in deciding ... how well a student has demonstrated his or her learning” (p 82). The syllabuses specify that performance standards are used to provide feedback to students and explain how they relate to the *assessment design criteria* and *specific features* used in designing tasks. Thus, the performance standards describe the

learning and provide the basis for task design, grading of student work and generating feedback to students.

Observations, teacher interviews and scrutiny of planning forum materials showed that teachers were introduced to the use of performance standards in task design, grading student work and providing feedback to students. Although the forums addressed the purposes of performance standards, in interviews almost a year later, participants could not recall being told their purposes. In spite of this, most were able, however, to accurately cite these purposes. In interviews, participants also identified diverse advantages and disadvantages of performance standards, although no evidence could be found for these in observation records, SACE Board documents or materials distributed at planning forums.

Observations, teacher interviews and scrutiny of materials distributed at subsequent processes focussed on *implementation* rather than *explanation* for the changes, with no evidence of reasons being given for the changes.

Overall, the evidence showed that teachers were well informed of the purposes of performance standards. There did not appear to be any attempt, however, to explain how performance standards represented an improvement on previous practice, nor identify disadvantages that might be associated with their use. The rationale for performance standards might be an area of assessment literacy addressed in the annual quality management processes for SBA, as discussed in section 8.4.

## **6.4 Learning opportunities in task design**

This section examines the quality management processes for evidence that they provided opportunities for teachers to develop skills in using the performance standards in designing assessment tasks.

### **6.4.1 Planning forums**

In the planning forums, performance standards were identified as an integral part of assessment and links were pointed out between the learning requirements, the assessment design criteria, the specific features and the performance standards. These connections were also demonstrated in the distributed material (SACE Board of South Australia, 2010e). The syllabus told teachers that tasks should “enable

---

students to demonstrate their skills, knowledge and understanding ... [and] allow students to demonstrate how they meet the performance standards” (SACE Board of South Australia, 2011d, p. 18). Neither the planning forum materials nor the syllabus explained, however, how the performance standards were to be used in designing assessment tasks.

#### **6.4.2 Clarifying support panel**

The two supervisor-participants described how clarifying support discussions improved their understanding. A flawed task chosen for the Physics clarifying forum had been selected because it was “a task that most teachers would know”, unaware that “it had some flaws in it”. The flaws only became apparent when the panellists examined student responses using performance standards. Although one student had correctly answered every question, higher order standards were not evident: there were no “open questions” that allowed students to demonstrate the “highest possible level of achievement” (SACE Board of South Australia, 2011d, p. 83). The supervisors had learnt that “previously-satisfactory tasks” were “too directed” when considered from the viewpoint of performance standards. Discussion also clarified terms used in the performance standards, such as “critically and logically analyses”.

Evidence indicated that supervisors had opportunities to learn about task design and performance standards from involvement in the clarifying support panel.

#### **6.4.3 Clarifying forums**

The materials distributed at the clarifying forums (SACE Board of South Australia, 2011i) make it clear that task design was addressed. A booklet, not subject-specific, had a checklist of “guiding questions for quality and suitability of an assessment task” (p. 6). The points were of a generic nature, for example, *Are the specifications from the syllabus considered?*

Interviews showed that participants were not satisfied with the clarifying forum regarding task design. Adam recalled the focus on assigning grades when he would have preferred more attention on task design, a desire expressed by five other participants. Frances was frustrated that the SACE Board had not been “willing to show test items or questions and identify which performance standard they are testing”. Russell believed that tasks needed to be designed with the performance

---

standards in mind, yet the tasks used at the clarifying forum “could’ve come from 1829”. He would also have preferred “stronger direction” on how many features should be assessed in each task. Alina suggested that “we didn’t get into ... task design” because of lack of time. Harry supported this, saying that hoping to deal with all of the changes in one clarifying forum had been unrealistic.

While teachers had opportunities to learn about the *importance of task design* in the clarifying forums, they had not learnt *how* to design appropriate tasks.

#### **6.4.4 Moderation panels**

The training provided for moderation panels included nothing that specifically addressed task design. Observation and interview comments, however, indicated that task design was frequently mentioned during discussion on students’ grades. All six moderator-participants agreed that the discussions had contributed to task design understanding.

#### **6.4.5 Exemplars**

The SACE Board exemplars were examined for evidence that they modelled commendable assessment practices and whether the exemplars were plentiful and updated regularly. The SACE Board web page on exemplars (SACE Board of South Australia, 2014b) indicated that exemplars were to “help teachers to design tasks that provide opportunities for students to demonstrate their learning” and that the annotations “describe what teachers may consider when developing assessment tasks and learning and assessment plans”. There was also a note that “Documents will continue to be uploaded in this section as they become available.”

Table 26 shows the Physics items for 2011 and 2014. Items on the same line are similar: some items have been renamed and some have minor amendments. Advice on writing a learning and assessment plan (LAP) was considered to be relevant and was examined, as was advice on writing assessment tasks.

**Table 26: SACE website materials 2011 and 2014: assessment design advice**

2011 exemplars	2014 exemplars
Annotated LAP exemplar	Annotated LAP exemplar
	Information sheet on LAP submission
	Stage 2 LAP checklist for teachers
Stage 2 Physics learning and assessment plan for use in 2011	Exemplar 1 - For a class with ready access to computers or laptops
	A brief guide to the Physics Issues Investigation
Circular Motion Practical Annotated	AT1 – Task 1
Issues Investigation Annotated	AT1 - Task 2
	AT2 - Task - Electricity and Magnetism test

The *Annotated LAP exemplar* states that its purpose is to assist all Stage 2 subject teachers (SACE Board of South Australia, 2013a). Consequently, annotations are generic and refer teachers to the relevant syllabus. The *Stage 2 LAP checklist for teachers* is based on the checklist provided to assist panellists. This, and the information sheet on LAP submission, is also generic. The 2014 Physics LAP exemplar was unchanged from 2011 and was used in approval panel training in 2011. It had no annotations to assist teachers.

*A brief guide to the Physics Issues Investigation* was based on material distributed at the clarifying forums in 2011. In 2014 there were three exemplar tasks without student responses (AT1 – Task 1, AT1 - Task 2, and AT2 - Task - Electricity and Magnetism test), one more than in 2011. Annotations identified the specific features assessed in each question, advised teachers on strategies for developing the task and informed them of the rationale used. In this way they provided advice to teachers constructing similar tasks. There was, however, no exemplar for the syllabus requirement that students design an experiment.

Ten participants appreciated the website exemplars, with Alina explaining that an exemplar “gives you something to work with”. Almost all participants modified exemplars before use, although Brenda confessed to having “virtually plagiarised” the Issues Investigation.

While participants made positive comments about the exemplars, they were also



critical. Colin had used exemplars to generate discussion among teachers and students about assessment expectations, but felt that the exemplars had left questions unanswered. Brenda would have liked exemplars with examples of questions that could be used to assess particular specific features; these would have been helpful in constructing tests and assignments. Frances found them “quite flawed” and had ceased looking for new exemplars. She was annoyed at the apparent contradiction of the SACE Board in “making us write tests to the performance standards [when] they’re not willing to show us how that is done”. Harry, having used the clarifying forum exemplar, had subsequently found it necessary to rewrite tasks to be more like his previous tasks and less like the exemplar.

Seven participants commented on the need for more exemplars, with four stating that a greater range of tasks was desirable. A comment about the shortage of exemplars had also been made on the Physics online forum. Colin initially said he would have liked more exemplars, but revised this, arguing that exemplars tend to stifle teacher creativity. Greg suggested that it would be instructive to have good and bad tasks with annotations explaining the differences, although others preferred exemplars that were “100% right”.

It appeared that the SACE Board website exemplars modelled commendable assessment practices, with annotations that offered relevant advice to teachers, but were insufficient in variety and number. They had been supplemented and updated to a minor extent, but they could not be described as plentiful.

#### **6.4.6 Feedback to teachers**

Written feedback was provided to teachers on their LAPs and, early in 2012, in the Physics Assessment Report (SACE Board of South Australia, 2012c) and school-specific moderation feedback.

##### **6.4.6.1 LAP feedback**

LAP feedback was restricted to comments regarding failure to complete the form correctly or comply with syllabus specifications, for example, failure to assess a specific feature or failure to have students design a practical investigation. Feedback on individual tasks was restricted to syllabus compliance and hence limited. One participant suggested that approval had focussed on “areas that are easy to administer”,

such as word count, rather than using the feedback to improve teachers' assessment strategies, which he believed were "the whole idea behind the assessment plan". Others described the reasons for non-approval as "trivial" and adjustments necessary for approval as "bureaucratic tweaks".

It was concluded that the feedback on LAPs offered no advice in the design of assessment tasks, but focussed on syllabus compliance.

#### 6.4.6.2 Moderation feedback

The form used by moderators for providing feedback had, for each assessment type, several tick-boxes, such as shown in Table 27.

**Table 27: Section of moderation feedback sheet**

<input type="checkbox"/> Evidence in the sample demonstrated achievement at lower grade levels. Particularly ...	<input type="checkbox"/> Investigation <input type="checkbox"/> Analysis & Evaluation <input type="checkbox"/> Application <input type="checkbox"/> Knowledge & Understanding
<input type="checkbox"/> task design limited the opportunity for students to demonstrate their learning in relation to the performance standards.	
<input type="checkbox"/> evidence presented departed from the syllabus specifications for ... <input type="checkbox"/> word count <input type="checkbox"/> no. of prac/issues tasks	

The moderators' form also had space for notes. SACE Board officers commented that "quite a few of the moderators" would have liked to provide more feedback in a "genuine attempt" to advise teachers of how to improve their tasks. They reported that moderators, looking from a teacher's viewpoint, said "I would want to know this". Adam described the limited feedback as "a real shame". David believed that it would have been helpful for teachers to be told that their tests had, for example, not included any extended response questions. Greg was one of three who would have liked to commend teachers where appropriate, but had been told "you cannot be specific". Kevin described the feedback as "terse and uninformative ... administrative gobbledegook" and reported the Board's view as "Our job's not to give feedback, our job is to assess." The feedback was "sanitised because it was

---

viewed to be too subjective”.

Schools were advised that they could contact the SACE Board for further information and Adam believed the SACE Board had been “hesitant to write anything down for fear it gets used against them”. This reflected Harry’s suspicion that some moderator comments might stimulate disagreement, a possibility the SACE Board wished to avoid.

The participants who had not moderated were also critical, describing the feedback as “minimal” and “inadequate”. One commented that, in the first year of the new SACE, feedback was particularly important for Physics teachers, without experience of social moderation. Thus task design advice in the moderation feedback to schools was considered inadequate.

#### **6.4.6.3 *Physics Assessment Report***

This was the first year in which the Physics assessment report (SACE Board of South Australia, 2012c) provided information about SBA and two of the ten pages addressed this. The following paragraphs are illustrative of the advice tendered and are not intended to be all-inclusive.

Much of the information had been available previously, for example, the reminder to “give students the opportunity to meet a range of performance standards at the higher levels” (p. 2). There were, however, clarifications of how opportunities might be provided, such as by “requiring the students to organise the data that they collect without teacher direction” (p. 2) and allowing students to suggest a range of improvements to an experiment, rather than restricting them to one. The report reminded teachers of the syllabus requirement for students to design and perform an experiment that tested an hypothesis.

Teachers were advised to assist moderators in their task by, for example, identifying the specific features addressed in each task. One comment advised that “Well-designed tasks ... guided students in their research and scaffolded how to reference” (p. 3). Test design advice included comments such as “A well-designed sequence of tests includes extended-response questions and experimental-skills questions” (p. 3), and advice that student skill development could be facilitated in a series of tests in

which extended-response questions increase in “length and importance” (p. 3) during the year. It was recommended that teachers provide students with summary sheets for assessment tasks.

One moderator-participant commented “some of the ideas ... perhaps weren’t written in as blunt a language as they could have been”. Another, also a moderator, said: “It didn’t tell me anything I didn’t know before”. While all participants intended to read the report, some had consciously left it until later in the year.

It was concluded that the report provided guidance on the design of assessment tasks but that some teachers may not have been aware that the advice related to their tasks.

#### **6.4.6.4 Advice on task design in feedback to teachers**

Although teachers received little task design advice in school feedback because of restrictions placed on moderators, the assessment report provided more specific advice. This was, however, of a general nature and not linked to the practices of individual schools.

#### **6.4.7 Communities of practice**

Teachers had opportunities to learn about task design through involvement in three types of communities of practice: the Physics online forum as well as intra- and inter-school collaborations.

##### **6.4.7.1 SACE Board online forum**

Of the 58 Physics online forum notices in the first year of the new SACE, those relating to SBA addressed a number of matters.

The first, in October 2010, sought to confirm the teacher’s syllabus interpretation regarding the number and form of assessment tasks. Six more followed on this topic, with some promoting splitting tests into two parts, thereby subverting the syllabus restriction on the number of tasks. This offered another example of failure to confront teachers’ misconceptions: the relationship between the quality of assessment and the number of tasks.

The first notice had also expressed concern about the capacity for students to receive support, such as from tutors, in SBA. This generated an exchange of ten notices

relating to ethics. One suggested a conflict of interest for teachers because of their role “to ensure that their students get the best possible mark”. The ethical implications in the term “best possible mark” were not confronted, either by other teachers or the SACE Board. The ensuing discussion addressed plagiarism, online purchase of assignments and advice on the detection of plagiarism. There was also a discussion on the practice of students re-sitting tests, one notice suggesting “a complete lack of guidance [from the SACE Board] in what constitutes fair and ethical assessment”. Within two days, this had been addressed by two SACE Board Executive Managers who stated: “It is not appropriate for students to repeat an assessment task ... once it has been undertaken” and “It is not appropriate to set summative tasks that are ... identical or very similar to the formative tasks.” They also referred teachers to the relevant policy (SACE Board of South Australia, 2012e). The author of the initial post thanked the officers and suggested the information should be more widely distributed.

Four notices commented on the performance standards, such as the meaning of the term “complex problem” (SACE Board of South Australia, 2011d, p. 84) and whether the specific feature I2, about selection and acknowledgment of sources (SACE Board of South Australia, 2011d), referred to a bibliography. Six notices discussed whether the expectation of students “designing and performing an experiment to test a hypothesis” (SACE Board of South Australia, 2011d, p. 80) meant that the designing and performing had to be done in the same task or whether they could be undertaken in separate tasks. The SACE Board offered no advice on complex problems, specific feature I2 or whether students were required to undertake the experiment they had designed.

Six notices discussed the possibility of teachers sharing tasks, but no tasks were posted. A first-year teacher sought an experienced teacher to provide a “2nd pair of eyes” for a test and received a positive response. Two notices discussed alternative formats of the performance standards, while another concerned weightings. One teacher, having received SACE advice that the method used for assigning levels and recording was satisfactory, offered “an example of one way it can be done”.

It appeared that the online forum in Physics informed teachers and allowed exchange of ideas, although sharing of tasks was limited. The SACE Board appeared to make

little attempt to stimulate discussion and generally limited its input to providing administrative information.

#### **6.4.7.2 Local communities**

The eight participants in schools with multiple Physics classes collaborated with other teachers in their schools. They shared common tasks, generally alternating their writing, although in one school the more experienced teacher designed all tasks.

Six participants were members of inter-school communities where tasks were shared and teachers “compared notes”. Teachers were also used as critics of other schools’ tasks. One group collaboratively analysed past exams in terms of the specific features, generating a resource used by all schools in the group. Another group contacted teachers with SACE Board experience for advice. This group also discussed the practical instructions provided to students, realising that excessive scaffolding would restrict students’ opportunities to achieve higher grades.

All participants working collaboratively expressed satisfaction, with task design the focus of most alliances.

#### **6.4.8 Summary of learning opportunities in task design**

For most teachers, education in task design was limited to clarifying forum discussions, website exemplars and the subject assessment report. Moderation feedback was considered inadequate by moderators and teachers. Teacher collaboration commonly involved assessment task design, with the greatest opportunities available to supervisors.

Teachers’ task design skills are an important link in establishing and maintaining effective SBA. As such, improvement in teachers’ task design skills should be an important goal of the annual quality management cycle for SBA and is addressed later.

### **6.5 Learning opportunities in grading**

This section scrutinises the quality management processes for evidence of learning opportunities in grading student work using performance standards.

### **6.5.1 Planning forum**

In the planning forums, teachers were presented with student work and asked to grade it using the performance standards, followed by discussion. Participants, however, commented that advice was lacking on *how* to undertake the grading, with Brenda describing the advice as “this is what you have to do, go away and do it”.

### **6.5.2 Clarifying support panel**

The supervisor-participants described how their opinions on grades assigned to student work changed through involvement in the panels. One reported how the panel, having chosen student work at an A- standard, by the time of the forum “thought it was a B+”. The other agreed to learning much from the experience: panellists became more familiar with the performance standards, and their application to student work, because of discussion among themselves.

It appeared, therefore, that membership of the clarifying support panel was instructive in using performance standards in grading student work.

### **6.5.3 Clarifying forum**

In preparation for the clarifying forums, teachers were asked to download assessment tasks and grade them prior to the forums. They were also invited to submit grades and comments via the website.

The SACE Board provided the grades submitted online by 28 teachers. Table 28 shows the frequency of teachers’ grades for each task (SACE Board of South Australia, 2010g) and the grades decided previously by the clarifying support panel.

**Table 28: Frequency of grades submitted by teachers online prior to clarifying forums**

Grade level	Frequency of grade level (n = 28)	
	Response 1	Response 2
A+		
A		1
A-		7
B+		10 = panel
B		8
B-		2
C+	1	
C	5	
C-	13 = panel	
D+	7	
D	2	
D-		
E+		
E		
E-		

The data showed that most teachers were within one of the panel's grade and that there were, marginally, more below the panel's than above.

Three participants used prior experience to determine the grades they submitted online, making no use of the performance standards. One said "it was so much quicker to just do it the old way", while a second described "a gut reaction from having marked thousands of pieces of work".

Participants' recollections of the forum discussions differed. Most recalled agreement among teachers and between teachers and the panel. Most thought agreement had been within one step on the SACE Board 15-point scale, A+, A, A down to E-. It is likely that different forums achieved different outcomes, but Table 28 data suggest similar grades were submitted online. All participants recalled teachers at the clarifying forums providing lower grades than the panel. Adam suggested that when Physics teachers gather, the outcome is that "testosterone comes at you" as each tries to prove "I'm really tough on my kids". It may be that in the public forum teachers were tempted to grade more harshly than in private or that the



harsher markers had not submitted their grades online.

At the clarifying forums, participants were aware of limitations in the training provided. Only one task was graded, not the full set of eight or nine required for moderation. Although teachers had been advised that grading with performance standards should not be applied to individual tasks, individual tasks were graded at the clarifying forums. The SACE Board attempted to counter criticism by saying the tasks were “indicative” of a particular grade. Participants considered that having limited samples of student work was due to lack of time, since each clarifying forum lasted three hours.

Presenters’ advice was reported differently by participants. John believed that the grading decision “made sense when it was pointed out”. Six others, however, felt that the grades were not well explained, and believed that few had been satisfied with answers to questions. Kevin believed “we were told what to think rather than ... thrashing it through and helping people adjust their mindset”.

From a presenter’s perspective, David reported how some teachers, reading parts of a student’s work, reached a decision about the student’s ability and wanted to apply that decision to other aspects without relevant evidence. He described some teachers as “very resistant to being told they’re not doing it right”, a view also expressed by Elena: “there’s a range of views and I don’t think many of them changed by the time they walked out”.

Practice in using performance standards in grading student work was undertaken at the clarifying forum. Although one participant described the forum as “useful”, it is suggested that the experience was of limited success with restricted time a contributory factor.

#### **6.5.4 Moderation panel**

Observations and interviews identified grading of student work as the major training provided to moderators. They were asked to grade three samples of student work that had been previously graded, followed by discussion that did not resolve all differences of opinion.

Moderators benefitted from hearing a range of opinions over the four-day moderation

period. Similarities between student work and the exemplars facilitated moderation. Differences between the work being moderated and the exemplars, however, necessitated extrapolation or interpolation, which generated moderator discomfort. Kevin was one who felt “under-prepared” for the moderation that followed and would have preferred a greater range of samples to be used. Peter would have liked more time to be spent on the training.

It was considered that, even though some would have liked more training, involvement in moderation had educated teachers in using performance standards to grade student work.

### **6.5.5 Feedback to teachers**

Feedback on the grading of student work was provided after moderation in advice sent to each school and the subject assessment report (SACE Board of South Australia, 2012c).

#### **6.5.5.1 Moderation feedback**

The moderation feedback used tick-boxes. Each assessment type had a heading: *Adjustments were made on the basis of...* followed by three text boxes. The text beside the boxes read:

- A+ sample(s) did not demonstrate sustained achievement at the upper A level. Particularly ...
- Evidence in the sample demonstrated achievement at higher grade levels. Particularly ...
- Evidence in the sample demonstrated achievement at lower grade levels. Particularly ...

The “Particularly ...” was adjacent to tick-boxes for the four assessment design criteria: Investigation, Analysis & Evaluation, Application and Knowledge & Understanding. Space was available for moderators’ notes. Moderators ticking boxes eventuated in feedback such as: Adjustments were made on the basis of evidence in the sample demonstrated achievement at lower grade levels. Particularly Analysis & Evaluation.

Moderators would have liked to give more feedback to teachers. Two participants,

for example, would have liked to advise teachers that their grades were “pretty close to being changed” and, as teachers, would like to have had such information. Kevin would have liked to make comments such as: “You had some A+ students who didn’t deserve to be there” but had been told moderators could not be so specific: “It had to be vague.”

The participants, as teachers, were also critical. Adam believed that teachers wanted “specific and detailed feedback as to why marks had changed”. Participants who had grades changed learned nothing from the feedback about the cause of the changes.

Thus feedback advice to teachers on the grading of student work was considered less effective than it might be. This is addressed in greater depth later.

#### **6.5.5.2 Physics Assessment Report**

The Physics 2011 assessment report (SACE Board of South Australia, 2012c) provided little advice to teachers regarding the use of performance standards in grading. What was provided was of an administrative nature, headed *OPERATIONAL ADVICE*. Nevertheless, the advice carried implications for teachers’ practices, for example, to provide moderators with “clear evidence” (p. 11) about how students’ grades had been determined and to provide “summary sheets for their students” (p. 11).

#### **6.5.6 Communities of practice**

It was possible that teachers might learn about performance standards in grading through the online forum and from local intra- or inter-school collaborations.

##### **6.5.6.1 SACE Board online forum**

Nine online forum postings related to grading. Two of these commented on the clarifying forum materials and a third was an extensive review of a clarifying forum just attended. This questioned whether the “significant agreement” among teachers at the forum would occur with “different tests set by different teachers” and questioned a moderation process “biased towards confirming grades”. It expressed concern at the lack of exemplars for practical tasks and another wanted exemplars demonstrating the full range of grades in a variety of task formats. One teacher asked: “Do I have to mark everything twice? Once as questions then as performance

standards?” No advice was forthcoming. Another notice offered a procedure for determining grades after marking student work.

It was considered that the notices raised questions rather than helping teachers to use performance standards in grading student work.

### **6.5.6.2 Local communities**

The eight participants in the schools with more than one Physics teacher all described practices ensuring equitable treatment of all students: all classes in each school used the same assessment tasks and procedures were implemented to ensure consistent marking between teachers. To establish consistent marking the first step, common to all schools, was writing a marks scheme that showed correct answers. Subsequent marking practices varied. The most common practice was for teachers to mark their own classes, followed by each teacher reviewing some work marked by the other teacher. In two schools, teachers took turns in marking all student work for a task; the other teacher then reviewed the marking before returning the work to students. In two schools, discussion between teachers was used to determine grade boundaries before work was returned to students.

Where participants collaborated with other schools, the process was much less formal. On one or two occasions marked tasks were shown to teachers from the other schools and discussion established that the standards appeared similar. One community reported discussions around interpretation of the performance standards: “What does it mean by *critically and logically* as opposed to *logically*?”

Participants who collaborated with other teachers, whether in the same or other schools, reported greater confidence in assessment as a result of discussions. The majority of Physics teachers who worked alone had limited opportunities for such learning. A means of creating collaborative opportunities for these teachers is offered later.

### **6.5.7 Exemplars**

SACE Board website exemplars were scrutinised for evidence that they might have provided reference points to assist teachers in making assessment judgements.

As in the earlier section on task design, the website was viewed in 2011 and 2014 for

advice to Physics teachers on the grading of student work. Table 29 shows the items identified, with similar items shown on the same line.

**Table 29: SACE website materials 2011 and 2014: grading of student work**

2011 exemplars	2014 exemplars
	Performance Standards - Stage 2 Physics
	AT1 – Task 1 – Student 1 Response
	AT1 – Task 1 – Student 2 Response
Issues Investigation A example	AT1 - Task 2 - Student 1 Response
Issues Investigation C example	AT1 - Task 2 - Student 2 Response

The most obvious feature of Table 29 is the increase from two items to five.

*Performance Standards - Stage 2 Physics* (SACE Board of South Australia, 2012i), added in 2012, had the adjectives used to discriminate between grades in red. Table 30 shows part of the document showing the grade descriptors for the specific feature, AE1: “Analysis of data and concepts and their connections, to formulate conclusions and make relevant predictions” (SACE Board of South Australia, 2014f, p. 79). In this table, the words coloured red in the original are in bold.

**Table 30: Performance standards for AE1**

Grade	Performance standard
A	<b>Critically</b> and <b>systematically</b> analyses data and their connections with concepts to formulate <b>logical</b> and <b>perceptive</b> conclusions and make <b>relevant</b> predictions.
B	<b>Clearly</b> and <b>logically</b> analyses data and their connections with concepts to formulate <b>consistent</b> conclusions and make <b>mostly relevant</b> predictions.
C	<b>Analyses</b> data and their connections with concepts to formulate <b>generally appropriate</b> conclusions and make <b>simple</b> predictions with <b>some relevance</b> .
D	<b>Describes basic</b> connections between some data and concepts and <b>attempts</b> to formulate <b>a</b> conclusion and make <b>a simple</b> prediction that <b>may be relevant</b> .
E	<b>Attempts</b> to connect data with concepts, formulate <b>a</b> conclusion and make <b>a prediction</b> .

It appeared the purpose of this document was to assist teachers in discriminating between grades when judging student work.

Table 29 includes two assessment tasks, each of which had two annotated, graded student responses. The annotations described the grading judgement for each

specific feature, with the final pages summarising the grades as highlighted performance standards. For each task, one response was of an A grade standard and the other a C grade.

Although the student responses were graded, the rationale behind the grading decisions might have been more fully explained. It may not have been clear, for example, why a task had been considered to show *critical and systematic analysis* of data (A grade) rather than *clear and logical analysis* (B grade) or why the student's conclusions were *logical and perceptive* rather than *consistent*.

In interviews, only Colin reported using exemplars for informing about standards but, rather than using them to inform himself, he used them to illustrate to students the quality of work needed for particular grades. The strategy generated difficulties, however, because he was not “100% clear” himself and was unable to answer questions satisfactorily.

It was concluded that the website exemplars only partially satisfied the attributes advocated in the literature. While they provided reference points to assist teachers in making assessment judgements, further explanations might have proved helpful. The exemplars had been supplemented, but it is doubtful they could be described as plentiful and they did not demonstrate each grade (Klenowski & Wyatt-Smith, 2010b, p. 114).

### **6.5.8 Summary of learning opportunities in grading**

For most teachers, education in the use of performance standards in grading student work was limited to the clarifying forum discussion. Although participants had found the exemplars useful in task design, they did not report similar appreciation of exemplars in the grading of student work. For a few, extensive grading practice occurred in moderation and clarifying support panels. Feedback provided to schools after moderation did not clearly explain why changes been made to teachers' grades.

Effective SBA is built on teacher competency in the consistent application of performance standards in judging student work. It is important that these skills be developed through ongoing experiences, such as the annual quality management processes, a matter discussed in the recommendations of Chapter 8.

## 6.6 Education re bias

Scrutiny of SACE Board documents, observation records, teacher interviews and officer focus groups revealed no explicit attempt to address the bias reported in SBA (Black, 1998; Crisp, 2010a; Harlen, 2005; Wyatt-Smith, 1999). As reported earlier, some teachers made grading decisions based on assumptions of a student's ability. There did not appear to be, however, recognition that basing grading decisions on what a teacher knows or thinks of a student constitutes bias. Consequently, no attempt was made to use this clarifying forum behaviour for professional development on bias in teacher assessment. One participant comment, however, suggested that some teachers may have gained some insight about the matter. Alina recalled presenters emphasising the need to grade work on what students had written and not factors such as handwriting.

Teacher bias is one of the major criticisms of SBA. Consequently, it is important that teachers are alerted to its existence and equipped to identify it and eliminate it. As discussed in Chapter 8, the annual quality management processes offer a means by which teachers may become better informed in this area.

## 6.7 Overview of the content of the quality management processes

The evidence showed that the content of the quality management processes had the potential to contribute to teacher learning. A strength of the processes was their using subject-specific strategies: meetings for teachers of one subject with most materials also subject-specific. The new SACE induction program introduced teachers to the use of performance standards, informing them of the changes in SACE assessment in 2011 and alerting them of the need to change their practices. The exemplars effectively informed teachers about assessment task design.

Although there were positive aspects to the processes, the potential for teacher learning was not realised as effectively as possible; teachers were not educated in the rationale for the changes and little advice was offered of how to make appropriate changes to assessment tasks. Advice and practice in the grading of student work were deemed insufficient and feedback was considered inadequate. Participants had commented on teachers at clarifying forums being more inclined to grade harshly,

but grades submitted online prior to the forums had not demonstrated a similar bias.

It is important that teacher learning in the use of performance standards, in task design and grading of student work, continues beyond the induction stage. It is recommended that the quality management processes be used to provide learning opportunities for all teachers. Strategies that might be used, such as website exemplars, improved feedback and greater teacher involvement in the annual cycle, have the potential to enhance teacher assessment literacy in the area of SBA and are discussed in section 8.4.

The next chapter examines the data for evidence of changes in teachers' beliefs and practices and whether these changes might be associated with involvement in the quality management processes.



## CHAPTER 7: RESULTS: TEACHER LEARNING

In this chapter, data are examined for evidence of teacher learning in the area of assessment that might be associated with involvement in the SACE Board's quality management processes.

The three sections of the chapter describe three forms of evidence: participants' recollections of their learning from the quality management processes, comparisons of participants' assessment beliefs and practices at the beginning and end of the research period, and examination of interviews to construct participants' learning journeys during the research period.

### 7.1 Teachers' descriptions of learnings from the quality management processes

Data on teachers' descriptions of their learning journeys included questionnaires and interviews, as well SACE Board exit surveys. This section analyses teachers' opinions on the assistance provided by the quality management processes in revising their assessment practices.

In each questionnaire, participants were asked about the most recent quality management process, followed by interviews about their experiences. Participants' approaches to learning are discussed in terms of whether they were of a surface, strategic or deep nature (Entwistle, 1998; Moon, 2000). Learning of a deeper nature is discussed with reference to Moon's (2000) stages of learning: *making meaning* refers to integrating new knowledge into the previous knowledge, *working with meaning* involves manipulation of knowledge for a specific purpose and *transformative learning* refers to participant evaluation of her or his frames of reference. Results for each quality management process are discussed in turn.

#### 7.1.1 Planning forums

Nine participants attended a planning forum and in the first questionnaire were asked to respond to four statements about the forums. Table 31 shows the statements and frequency of responses.

**Table 31: Participants' opinions on helpfulness of the planning forums (n = 9)**

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
The planning forum helped me to write an approved learning and assessment plan	1	2	0	6	0
The planning forum helped me to design assessment tasks	3	0	0	6	0
The planning forum helped me to use the performance standards in marking student work	1	1	1	6	0
The planning forum helped me to apply the performance standards in a manner consistent with other teachers	2	2	1	4	0

A majority (67%) agreed that the forums were helpful in writing a learning and assessment plan (LAP), designing assessment tasks and using the performance standards in marking. As a group, they were neutral about whether the forums had helped them to develop consistency with other teachers in using the performance standards. The interviews generally supported these questionnaire responses, although five reported leaving the forums with some uncertainty about aspects of the new assessment system.

After the planning forums, the SACE Board invited teachers to complete exit surveys on confidence with the new assessment system. Data from the Physics planning forums in Table 32 show the percentage distribution of responses.

**Table 32: Physics teacher assessment confidence (percentages) from planning forum exit surveys (n = 101)**

Item	No response	Not confident	Slightly confident	Confident	Very confident
Confidence in developing familiarity with the subject outline, particularly student capabilities	0	7	30	59	3
Confidence in understanding the purposes of Quality Assurance (fairness, validity, reliability)	3	12	29	42	14
Confidence in dealing with insufficient evidence of student learning	0	10	39	43	7
Confidence in recording and combining results at the assessment type level	3	12	36	45	4
Confidence in understanding the processes of the Quality Assurance Cycle	1	17	33	42	6
Confidence in using the performance standards to assist in the design of assessment task	0	7	46	45	1
Confidence in using the performance standards to assess evidence of student learning	0	12	42	46	0
Confidence in designing or reviewing an assessment task using the assessment design criteria	0	10	45	45	0
Confidence in using the performance standards to support student learning	1	12	46	41	0

The responses show that, apart from the first three items, less than 50% of the teachers rated themselves as “confident” in relation to being able to implement the new assessment system. Teachers were invited to respond to the question “What further clarification could assist you to teach Stage 2 in 2011?” (SACE Board of South Australia, 2010c). Overall, 39 Physics teachers (39% of attendees) responded. Most responses identified factors limiting the forum’s educative value, including the need for more exemplars (mentioned by ten), brevity of the forum (six) and unclear wording of performance standards (five). Six mentioned the need for further assistance in using performance standards to grade student work and four wanted more assistance with task design.

The questionnaire items (Table 31) suggested that the forums were “helpful” in learning about key aspects of the new assessment system. Teacher responses to the

items relating to knowledge about the assessment changes (Table 32) appear to support that the forums were “helpful”. The exit surveys, however, showed substantial numbers were not confident in implementing the new system and the open-ended questions identified areas of teacher uncertainty. It is clearly possible that teachers could report the forums as “helpful”, but still not be confident about implementing the changes.

In interviews, participants were asked open-ended questions about the planning forums. Most described them as a useful introduction to the changes and the need to work with “rubrics”. Most were comfortable that existing assessment tasks could be adapted to comply with the new requirements and did little preparation in the eight months before 2011. Kevin and the other Physics teacher in the school had, however, “decided what we were going to do and how we were going to do it” by the end of 2010. John had been alerted, by the tasks used in the forum, that adapting previous assessment items might not be appropriate. Adam saw the requirements of LAP approval and moderation of student work as restricting the number of tasks, with implications for teachers who used assessment as “a form of classroom management”. For Adam the planning forum raised the necessity to discriminate between formative and summative tasks, leading to deeper learning through reflection on the school’s assessment practices.

Thus, at a surface level, the planning forums were helpful but did not appear to engage deeper learning; this occurred later, when teachers were required to work with the new assessment system in preparing assessment tasks and grading student work. It appeared that teachers believed that the planning forums contributed to learning about the new system, although, as might be expected, since the learning was of a surface nature, they were not yet confident of the capacity to implement the changes.

### **7.1.2 Supervisors**

Supervisors were members of LAP approval, clarifying support and moderation panels and experienced more than 50 hours of involvement in the quality management processes during 2011, as reported earlier.

Two participants served as supervisors in different subjects. Both described an

evolutionary process from which they had learnt “quite a lot”. One described how LAP approval had alerted him to different assessment sequences and novel tasks, both of which caused him to review his own assessment practices. Both described how, through discussion, they became aware of flaws in tasks selected for the clarifying forum and how panel discussion had clarified understanding of the performance standards, essential for their roles as clarifying forum presenters. It was decided in discussions, for example, that *problem-solving* required more than substitution in mathematical formulae; this was determined to constitute *using formulae*. They also came to realise that tasks should not be “too directed” but needed to allow better students to demonstrate their understandings. Supervisors were, it appeared, beginning to move into the area of deeper learning. They were required to make meaning of the performance standards and work with meaning in order to lead teachers in the clarifying forums.

One participant was grateful for the cooperative contact with other supervisors to “make sure that what I was thinking matched up with what people I trust were thinking”. The other commented on the importance of coming to understand the SACE Board officer’s thoughts and reporting teachers’ thoughts to the officer. Both described an evolution of understanding that developed in clarifying forum preparation and continued beyond the forums. Both described an “iterative process where shared understandings grow over time” (Hipkins & Robertson, 2011, p. 19) that appears to fit the aspirations of the SACE Review (Crafter et al., 2006).

It appeared that ongoing discussion about the use of performance standards, both in task designing and grading student work, contributed to teacher learning. This apparent demonstration of the potential for teacher learning in quality management processes stimulated considerations of how this potential might be realised for more teachers.

### **7.1.3 Clarifying forums**

Participants’ experiences of the clarifying forums were examined in items in the second questionnaire and in the corresponding interviews. Questionnaire items asked about agreement with a number of statements, with responses shown in Table 33.

**Table 33: Participant responses to the clarifying forums (n = 13)**

	No response	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
The clarifying forum helped me to design assessment tasks	0	1	2	3	7	0
The clarifying forum helped me to use the performance standards in marking student work	0	2	3	1	7	0
The clarifying forum helped me to apply the performance standards in a manner consistent with other teachers	0	1	3	4	5	0

Seven participants agreed that the clarifying forum had helped them to design assessment tasks. In subsequent interviews participants had the opportunity to expand on the reasons for their opinions. The flawed task mentioned earlier was recalled by participants as “too prescriptive”, with Elena explaining “you can’t give them everything because you need to let them introduce the complexity”. Harry reported, however, that “how to make it more open, wasn’t really done” and Adam reported teacher frustration because “there really wasn’t any explanation of how to write the tasks”. Greg was unclear how to balance being “very specific about what students are supposed to produce” and “over-specifying”, possibly jeopardising student results. His dilemma remained unresolved at the end of the research period. Clarifying forum observations confirmed that no advice was provided on *how* to design appropriate assessment tasks.

The forum raised participant awareness of the importance of assessment task design, with a number subsequently revising their tasks: Adam began including the performance standards with each task, Elena and Peter began recording students’ manipulative skills and Greg had students write about their group work skills. Peter, whose students had recently begun the issues investigation, withdrew the task and issued a revised version because he became aware of the need to have the students formulate a question. Others ensured inclusion of more questions “to discriminate between the very top students”.

It was concluded that the clarifying forum had “helped” by alerting participants to the importance of designing appropriate tasks, but that insufficient advice had been

provided about their construction. Thus learning in task design appeared restricted to a surface approach, with little opportunity for deep learning.

Clarifying forum assistance in improving consistency in performance standards application was reported by five participants in questionnaires (Table 33). In interviews, participants disagreed on the extent of consistency, possibly because they had attended different sessions. Adam, for example, recalled most teachers being “pretty close to the mark” whereas Frances recalled a “huge range of variation”. Reports of discussions also varied, with most participants reporting inadequate explanation for the grades assigned by presenters. Russell had been “a little stunned” when, on asking why a piece of student work was of a particular grade, had been told “because it is.” On the other hand, John believed “it certainly made sense when it was pointed out”. Two factors were mentioned as confounding any movement towards consistency between teachers: teachers resistance “to being told they’re not doing it right” and the limited time for discussion. Colin, on returning to school, reviewed previous grades, having learnt that he had been too harsh.

Although a number of participants agreed that the clarifying forums had “helped” in grading, narrowing the range of opinions somewhat, most thought the narrowing was insufficient. Clarification of the performance standards, such as the above reference to “problem solving” (see 7.1.2), had assisted teachers. Teacher resistance and limited time, however, appeared to have restricted the opportunity to resolve grading differences. The differences between the learning of supervisors and that of teachers suggested that changes to quality management processes might improve their educative value. Recommendations that might enhance the learning of more teachers are discussed later (see section 8.4).

#### **7.1.4 Moderation panels**

It was anticipated that moderator-participants might provide evidence supporting the reported educative value of social moderation (ARG, 2006; Harlen, 2004a; Matters, 2006), cited in the SACE review (Crafter et al., 2006) as a reason for recommending the adoption of social moderation. In pre-moderation interviews, most participants anticipated moderation as being potentially educative. Post-moderation questionnaires and interviews were examined for evidence of the benefit of the experience for those who served as moderators. Table 34 shows the questionnaire

responses.

**Table 34: Moderators' opinions of the educative benefit of moderation (n = 6)**

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Serving as a moderator will help me to write an approved learning and assessment plan	0	1	1	3	1
Serving as a moderator will help me to design assessment tasks	0	0	0	3	3
Serving as a moderator will help me to use the performance standards in marking student work	0	1	0	3	1
Serving as a moderator will help me to apply the performance standards in a manner consistent with other teachers	0	0	1	4	1

The data indicated that most moderator-participants agreed that moderation had helped with their assessment practices, more so with task design.

All with moderation experience, expanding on these in interviews, acknowledged the educative value of being a moderator, although scrutiny of transcripts showed variety in the depth of learning that had occurred.

The most obvious learning was of a surface-level nature; for example, regarding the administration of moderation, that moderators worked in pairs and that grade changes required agreement of three moderators. Adam believed that moderation had been more useful than the clarifying forum because moderators saw more samples of student work and had more time for discussions. Greg also mentioned the time and commented on the benefit of discussions with a variety of moderators. All six moderator participants reported the benefit of seeing other teachers' practices. Some participants gained confidence in their own practices from seeing poor tasks that were "unduly repetitive" or that provided excessive scaffolding.

From their moderation experiences, a number learned strategic approaches for future submissions. Adam described assessing "to give the kids the best opportunity of being successful rather than really assessing their knowledge ... to make sure that it gets reflected in the best light possible". David was going to add "coloured paper"



notes to assist them. Greg and Kevin planned to redesign tasks to be “moderator-friendly”, for example, by “putting the specific features in places that are easy for the moderator to find.” A strategic approach, however, was not restricted to learning from moderation: the term “playing the game” had been previously used by teachers and SACE Board officers. Colin used the term to refer to practices that improved grades without improving student learning: “learning [and] good grades ... are not always the same thing”. Greg suggested that assessment had become “a game of producing polished-looking work to impress moderators.” A number of the participants described actions they took to benefit their students. Kevin only wrote positive feedback on student work; negative comments were made orally. Peter had adjusted the timing of his tasks, with all tests completed in the last month before moderation. Although participants mentioned strategies learned from moderation, it appeared that strategic practices preceded the new SACE. There were few acknowledgements that the previous system was subject to manipulation, for example by failing to do practical work. Most seemed oblivious to “playing the game” in the previous system.

In the training provided, moderators were required to work with meaning: they were given exemplars that they graded and discussed. Although five believed that consistency had been achieved, three would have liked more examples and longer discussion. The sixth felt that revealing in-house discussions would breach the confidentiality agreement signed by all moderators. One participant identified a procedural flaw: training involved individual pieces of student work whereas moderation involved folios of four or five tasks. Most agreed with grades assigned. David, however, suspected that some moderators “didn’t agree ... and then didn’t work at that standard”. Adam opined that a few were frustrated because the reasons for some grades “never seemed to get resolved [so moderators had to] live with it”.

Working with meaning continued when moderators used their experiences to modify their own assessment practices: Adam reduced the number of tests and Greg omitted one practical and had students undertake fewer tasks under direct supervision. Adam had also learned the need for more attention to students’ interpretations of experimental data, causing him to reflect on the “recipe-like” experiments he had used.

Moderation experience also prompted reflection on the fairness of the process. Two participants commented on the need for agreement between two moderators and a supervisor for changes to be made. On the other hand, five raised concern over the “arbitrariness” of the process, with time constraints a concern. Three suggested different outcomes might occur from moderator fatigue over five days. Kevin suggested that “recalibration” might be of benefit after a few days. Peter was one who believed that grade changes were sometimes less than desirable because differences in task complexity and teacher support (e.g., scaffolding, oral advice) were not sufficiently acknowledged. SACE Board officers acknowledged that moderation outcomes varied and cited an instance where, accidentally, a moderator had been asked to re-moderate the same school, with a different outcome.

Because some of the participants had previous moderation experience, reflection on the role and fairness of social moderation was not restricted to moderator-participants. Colin explained how moderator discussion “over time [generated] ... a greater collective understanding” of the performance standards (Hipkins, 2010b; Hipkins & Robertson, 2011; Klenowski & Wyatt-Smith, 2010b) and contrasted this with the absence of this with statistical moderation. The terms “secret business” and moderators being “in the know” were used about moderators being privy to information denied to others. One participant mentioned that, after a few years, the International Baccalaureate (IB) had begun advising all teachers of instructions provided to moderators and suggested that something similar might happen in the SACE.

The most common learning reported by moderator-participants related to the benefits of seeing a range of student work, which usually enhanced confidence in their own practices. The opportunity for days of discussion was also reported as very beneficial. A common learning was of strategies that might be used to benefit their students in the future, such as ensuring that work was presented in a manner that made the moderation task easier. The majority had observed practices that did not engender confidence in social moderation, particularly its lack of consistency—both between moderators and over time—and failure to account for different assessment conditions in different schools.

Although consistency between moderators in grading student work was reported, the

most common reported learnings were increased awareness of appropriate assessment tasks, increased confidence in one's own assessment and improved presentation that might benefit students. Most of the learning was of a surface-level nature, but awareness of a strategic approach was raised and some were stimulated to think more deeply about the implications of social moderation. Similar observations were noted with many participants: although quality management processes operated at a surface level, teachers' subsequent reflections were in greater depth. The quality management processes served as stimuli for deeper learning. Deeper learning—reflection, making meaning, working with meaning and transformative learning—occurred as teachers attempted to resolve issues that arose from the new learning. It is suggested that depth of thinking might have been associated with personal characteristics of the individuals—“how they perceive themselves, how they present themselves, what they consider important” (van den Berg, 2002, p. 582)—rather than their roles, locations or sectors. Collection of such personal data was not attempted in the current research.

### **7.1.5 Feedback after moderation**

For non-moderators, the opportunity for learning from moderation was restricted to the feedback provided in two ways: school-specific information and the Physics assessment report (SACE Board of South Australia, 2012c).

#### ***7.1.5.1 Individual schools' moderation feedback***

All participants had read their school-specific feedback and expressed disappointment in the lack of information provided. The feedback invited teachers to contact SACE Board officers for further information, although few accepted this offer. Moderator-participants did not do so because “I knew there'd be nothing there”. The principals of three participants' schools had instructed all teachers to contact the SACE Board for further information. Russell reported the calls yielded “no explanation at all”, while John described the outcome as “pretty poor”.

Participants would have liked more feedback, even when no changes had been made to grades. Adam described his frustration: changed grades showed something was wrong but the feedback did not identify the problem. Colin had grades changed but, after reading the feedback, was “none the wiser ... of what I should do to change”.

Russell described his school as “fortunate” to have a moderator on the staff, who assisted other teachers to interpret the feedback provided. He described this contribution as “brilliant”, but was critical of the need to have “inside knowledge” to interpret feedback.

That most teachers failed to contact the SACE Board raises the question of why they failed to do so. Three possibilities were offered by participants: moderator awareness, historical awareness and lack of concern. Moderators awareness has been mentioned and five participants reported, either from personal or other teachers’ experiences, that previous SACE feedback had also been deficient. Some participants failed to contact the SACE Board because the changes were considered minor, for example, one student had a C changed to a C-.

Overall, the interviews indicated that moderation feedback was not specific enough to inform teachers of how to improve their assessment tasks or their grading practices.

#### **7.1.5.2 Physics assessment report**

Given that the school-specific feedback had not been as informative as hoped, participants anticipated that they might learn from the Physics assessment report. At the conclusion of the research period six participants had read the report, with the others intending to read it later. Two believed it was more informative than previously. Russell believed the reports in most subjects were an improvement on previous years’ because they identified assessment design criteria that had consistently been addressed poorly in a subject.

Most believed the report could have been more specific. Adam believed that it accurately reflected moderators’ observations, but that the comments “perhaps weren’t written in as blunt a language as they could have been”. Greg, who had moderated, reported that it was more meaningful because of his involvement in moderation, although he learned nothing new. Colin hoped to find an explanation for the changes to his students’ grade, but “there was nothing there that gave me a clue”. He described how the equivalent IB information had become more specific recently and hoped that this might happen with the SACE. Two participants had contributed to writing subject assessment reports. One had attempted to provide feedback for

individual schools but that, given the format of the report, had found this difficult.

It appeared that the subject assessment report had improved on previous ones, but that it still suffered from a lack of specificity, making it difficult for teachers to link the information to their own practices.

### **7.1.6 Exemplars**

This section examines the use that participants had made of exemplars in their school-based assessment (SBA) during 2011 and their contribution to teacher learning.

Twelve participants used the exemplars. Although Greg reported they were “by the far the most important things” the SACE Board had provided, he was one of eight who criticised their insufficient number. Adam reported a SACE Board manager saying that it was the teachers’, rather than the SACE Board’s responsibility, to provide teacher materials. Adam agreed that teachers had this responsibility, but that there was a need for SACE Board support, particularly while teachers were unfamiliar with the new requirements.

Four participants criticised the exemplars’ quality. Elena believed the exemplar practical was too simple: its questions were so closed that achieving an A grade was difficult because they did not allow students to “introduce complexity” in their responses. Harry had modified his practicals to match the exemplar but was frustrated at the clarifying forum, “where they said the exemplar is rubbish”. Brenda wanted exemplars that showed questions addressing the different specific features. Frances was frustrated with exemplars that did not assign specific features to questions when schools were expected to do so. In one observed clarifying forum, there was much heated discussion on the failure of the SACE Board to provide this information. Although SACE Board officers were aware of teachers’ reactions, they had been instructed that no exemplars were to show specific features assigned to questions.

Participants used the exemplars in various ways. None used them without some modification. Brenda described adapting one slightly “to make sense in my head”. Kevin “didn’t like” them but had used parts, as had most participants, as models for

his assessment tasks. Colin and Peter used them in student discussions, to show the standard required and how grades had been assigned against the specific features.

Overall, the interviews suggested that the participants used the SACE Board exemplars to enhance their knowledge and practice, but would have liked more. It appeared, however, that learning from the exemplars would have been enhanced with more exemplars and more guidance on designing tasks to address each specific feature.

### **7.1.7 Communities of practice**

This section examines interview data for evidence of the extent, nature and benefit of collaboration between the participants and other Physics teachers. No participant posted a notice on the Physics online forum during the research period and none mentioned it in the interviews, but other forms of collaboration were mentioned by participants.

Three participants effectively tackled the changes of 2011 alone. Discussions with other Physics teachers had been limited to rare social meetings, insufficient to comply with the concept of a community of practice. Brenda described getting “more and more depressed” at having to work alone, with no “offsider to bounce off”. Peter described how he had “fought my way through it”. This section examines the collaborations of the other ten participants.

#### ***7.1.7.1 Intra-school collaboration***

Seven participants with other Physics teachers in the same school reported their collaborations as continuing previous practice. For Elena, the collaboration was new; she and the other Physics teacher were both new to the school.

The eight described different levels of collaboration. Two faculty heads provided leadership as previously, with little input from the other teachers. The other six described the benefit of having a second opinion in task design and grading student work; discussion allowed them to clarify and revise their thoughts. Adam met with the others to discuss how to design tasks that addressed the performance standards. They also cross-marked to confirm that “an A in one class is the same as an A in another class”. Two teachers had tried assigning specific features to questions in a

test. When they generated different assignments they felt no need to resolve the differences because they believed they could plan good tests without using specific features; they dismissed the assigning of specific features as “generating work that doesn’t need to be generated”.

Elena explained how the two had shared task design; one with a background in practical physics wrote the practical tasks while the other wrote the tests. Initially, each marked the work of half the students; they would then swap to confirm the other’s marking. Later, satisfied they were consistent, they took turns to mark all responses for each task. Although consistent with each other, they had no external benchmark: “We just have each other.” She described the relationship between two teachers with very different prior experiences as “very helpful”.

Kevin reported how, for each task, one teacher would set the task and the other would mark it. They had agreed on the need to include a range of questions and ensure that some required higher order thinking, without needing to refer to the specific features. Marking would be checked by the other teacher, with discussion where different marks might have been assigned. After they had agreed on marks, work was returned to students.

John reported that the teachers had agreed to use common assessment tasks and to cross-mark. He valued the ready access to another teacher to “work through the issues”, making the new system “much easier”. Colin and colleague had also used common assessment tasks and cross-marked; they were perplexed when one class had changes made at moderation but not the other. As already mentioned, feedback did not resolve the dilemma. They intended to follow up with the SACE officers but had not done so by the end of the research period.

Apart from two faculty heads, participants described ongoing discussions with others as supporting their learning and mastery of the new assessment system, reflecting the literature reports of the benefit of collaboration to teacher learning (Hawley & Valli, 1999; Timperley et al., 2007; Wei, Darling-Hammond, & Adamson, 2010). Overall, the sustained collaboration was an important element in teacher learning, with participants more confident than others who had worked in isolation.

### **7.1.7.2 Inter-school collaboration**

Six participants collaborated with teachers from other schools, two as subject supervisors as described previously. The following section describes participants' other inter-school collaborations, all with teachers from the same sector.

Three were members of a group that met approximately twice each term during 2011. Confidence in the other members arose from prior contact as members of SACE marking panels. Colin described the practice as “comparing notes”: sharing observations of forums and discussing consistency of information provided at different meetings. The group went through the previous SACE Board examinations and assigned specific features to each question. While ideas were exchanged, each school operated independently: “confirming that we’re on the same track”, while “agreeing to differ”. Although most of the time had been spent on task design, they also examined student work to “establish benchmarks”. Meetings became less frequent over the year and, in 2012, emails were exchanged but there had been no meeting by April, when the last interview was conducted. One participant reported that two members had moderated and were willing to share their experiences. Communication to the facilitator had elicited no response, which puzzled the participant: “If I was them I’d be hungry for information.”

Greg described a group of five teachers who met “two or three times” and exchanged information by email. Again, the members had known each other previously. Greg described the meetings as “the most important event over the last two years, .. more useful than the clarification forums” because the teachers exchanged resources. To one meeting they invited another teacher with extensive SACE Board experience, known to Greg through SACE Board marking experience. Greg also contacted another teacher, known in the same way, to critique his assessment tasks. Greg, aware that his access to these teachers was through shared panel experience, realised that similar access was not available to other teachers and was critical of the SACE Board for not providing it to all.

Harry was a member of a group of approximately ten schools that met once. Each teacher brought some practical tasks that were viewed and discussed to ensure they were marking to similar standards.



All participants attributed some learning to these inter-school collaborations, primarily clarification of their understandings through discussion with other teachers. Participants also gained confidence from confirmation that “we’re on the same track”.

### **7.1.7.3 A summary of collaborative practices**

Access to ongoing discussion was commonly described as a positive outcome of the collaborations, leading to greater understanding. All participants involved in both intra- and inter-school collaboration, however, reported the intra-school collaboration the more valuable. This may have been due to its greater frequency or because discussions were embedded in the teachers’ work (Wei et al., 2009).

Recommendations to extend the opportunities for teacher collaboration are discussed in section 8.4.

### **7.1.8 A summary of participants’ perceptions of learning from the quality management processes**

Overall, participants believed that the quality management processes had contributed to knowledge development of performance standards and skills in their use. The processes had stimulated learning and provided useful resources. Participants were, however, critical of a number of aspects: insufficient time to resolve differences, the limited number and quality of exemplars and limited post-moderation feedback. In most instances, the learning from the quality management processes was of a surface nature. The exceptions were the supervisors who had access to more time and some moderators who reflected on their observations. The opportunities for deeper learning occurred after the initial induction program. It is recommended that induction programs be extended to facilitate learning beyond the introduction of a new program, as is discussed in section 8.5. The data also indicated that the quality management cycle has the potential for effective teacher learning, but that modifications would be needed to enhance their educative value, as discussed in section 8.4.

## **7.2 Evidence of teacher change**

The current research examined participant awareness of the reasons for the changes introduced in the assessment system, particularly the use of performance standards.

The research sought evidence of changes in assessment beliefs and practices, which might be linked to involvement in the quality management processes.

### **7.2.1 Teacher awareness of the reasons for change**

Adults are more motivated to learn if convinced of its benefit (Donavant, 2009; Fogarty & Pete, 2004; Knowles et al., 2005). Consequently, teachers might be more willing to change if a coherent rationale had been provided for the new assessment system. Conversely, willingness to change might be hindered without sufficient rationale for the changes. Hence teacher awareness of a clear rationale for the changes was explored.

Teachers were introduced to the new assessment system in the planning forums, and it was there that a rationale for the changes might have been expected. The planning forums had been conducted prior to the current research and the first interviews explored teacher awareness of the reasons for the changes. Participant knowledge of the reasons appeared sparse, with Kevin describing “change for the sake of change” being “foisted” on teachers who saw no need for change. No participant could recall reasons being given at the planning forums, although one recalled presenters “reading the party line, ... not necessarily convinced that they believed it”. Two believed they had been told it was introducing “a better way of assessing” but why it was better was unexplained. Observation of planning forums and scrutiny of materials revealed a focus on requirements of the new system rather than rationale behind the changes.

When asked the reasons for the change, one participant confidently cited a “push ... [for] more uniform ... assessment processes” identified in the SACE Review (Crafter et al., 2006), which was also tentatively suggested by another two. Others suggested a range of reasons. Five believed that the changes were politically motivated to improve SACE completion rates. Two saw the change as a “shift towards the humanities”. Four, aware of the ongoing work on the National Curriculum, considered the changes an interim measure. Three suggested they had been introduced to stimulate change in teachers’ practices and one suggested that the changes had been introduced “to get rid of the old wood” among teachers. Another proposed that it was to promote the use of *assessment for learning*.

From participant failure to recall being told the reasons and the diverse reasons proffered, it could be argued that the planning forums failed to inform teachers adequately of the reasons for the changes. Consequently, the changes might be met with resentment and a lack of motivation (Donavant, 2009).

## **7.2.2 Changes in teachers' assessment beliefs**

First and third round questionnaires and interviews were used to explore possible changes in participants' beliefs in three areas: the use of SBA, performance standards and social moderation.

### **7.2.2.1 Beliefs regarding school-based assessment**

Questionnaire responses were used to generate scores for two derived variables relating to beliefs in SBA: *confidence in formative* and *confidence in summative*, although having only thirteen participants was recognised as a serious impediment to data quality. When Rounds 1 and 3 data were compared, participants commonly gave identical responses in both rounds, with occasional differences of one on the 5-point scale, even though the two rounds were completed almost a year apart. Comparisons of Rounds 1 and 3 data for other derived variables yielded similar results. Because of the lack of change, it was decided these quantitative longitudinal data would not be examined further. Possible explanations for the stability in the data are examined later.

In interviews, participants were asked open-ended questions about reasons for the summative use of SBA in the SACE. Responses in the two rounds were compared and found to be very similar, with small upward and downward changes in frequency of different reasons. The two most common reasons for the summative use of SBA were the inability of examinations to assess some skills (given by seven participants in Round 1 and four in Round 3), the advantage of assessing skills more often (six in Round 1 and two in Round 3) and the detrimental effects of examination conditions on some students (two in Round 1 and five in Round 3). For all reported advantages there were, however, inconsistencies. For example, the less stressful nature of SBA was mentioned by different participants in the two rounds, but nobody mentioned it in both rounds. The inconsistencies render any conclusion problematic. A lower frequency in Round 3 might suggest that the advantage had been mentioned in the

planning forums and forgotten over the research period. A higher frequency might suggest that the advantage had been mentioned during the research period. That mixture of increases and decreases suggested that the frequencies were unrelated, in any systematic way, to involvement in the quality management processes.

Participants were also asked about disadvantages of making summative use of SBA. Again, comparison of the two rounds was undertaken. The major disadvantages identified were the inconsistency between schools (mentioned by 9 participants in Round 1 and 5 in Round 3), the openness of SBA to manipulation (1 participant in Round 1 and 6 in Round 3) and the inability to verify student work (3 in each round). Again there was a mixture of increases and decreases with little change overall (15 disadvantages mentioned in Round 1 and 17 in Round 3) and inconsistencies; inability to verify student work was mentioned by different participants in the two rounds. The most apparent change was the large increase (from 1 to 6) in describing SBA as open to manipulation. This increase was examined in more detail. In Round 3, five of the six moderators identified the possibility of SBA manipulation being problematic and scrutiny of the interviews suggested that the comments might have been related to moderation experience, involving a mixture of rumour, misunderstanding and evidence. Adam, for example, reported “you hear all kinds of stories” of formative tests identical to the summative test. Kevin described teachers who graded weaker students with better, “hoping that they would get through”, but acknowledged the possibility that the teachers “genuinely thought they were A+ students ... [even though] there was daylight between them”. David reported classes where all had “the same conclusion, ... the same set of data, ... the same slope for their line of best fit”.

Overall, there was little evidence of change in teacher beliefs regarding the use of SBA, which might be associated with involvement in the quality management processes, with the possible exception that moderation raised awareness of the possibility of SBA being open to manipulation.

### **7.2.2.2 Beliefs regarding performance standards**

This section examines evidence for possible change in participants’ beliefs regarding the purposes of performance standards, their advantages and disadvantages, which might be related to involvement in the quality management processes.

Participants' opinions of the purposes of performance standards were collected by questionnaires and interviews in Rounds 1 and 3. In the questionnaires, participants indicated agreement with statements that performance standards were important in task design and the grading of student work. There was a small decrease in agreement that performance standards were important in task design: eleven in Round 1 and nine in Round 3. Responses about the possibility of grading without performance standards were similar in the two rounds: ten and eleven. It was considered that these differences were insufficient to indicate change in participant beliefs.

Interview comparison was based on the twelve participants who undertook interviews in Rounds 1 and 3. In Round 1 all twelve cited their use in designing assessment tasks, changing to eleven in Round 3. For grading student work, the numbers increased from six to ten. In both rounds, only John stated that the performance standards helped teachers to provide useful feedback to students, but believed they were more appropriate in the middle school.

Because the use of performance standards in task design, grading student work and providing student feedback had been mentioned in planning and clarifying forums and in SACE Board documents (SACE Board of South Australia, 2010e, 2011d), it appeared that participant knowledge of these purposes might be due to involvement in the quality management processes. Three participants mentioned in Round 1 that the performance standards had made them think about what was important in the subject. Observations of forums and scrutiny of SACE Board documents found no reference to this purpose. Hence, it was considered that these participants were evaluating performance standards on the basis of their own experiences, transformative learning, rather than on the basis of the quality management processes.

The advantages of performance standards were addressed only in the interviews. In Round 1, ten advantages were mentioned by six participants, whereas in Round 3, sixteen were mentioned by ten: an increase in the number of advantages and the number of participants. Mention of their assisting task design grew from four to nine. Brenda, for example, described performance standards as "extending to the next stage" the use of criteria in the previous syllabus (SACE Board of South

Australia, 2010). She was “using the questions I’ve always used, ... just perhaps tweaking them a bit”. Colin described performance standards as “an extra checking mechanism” in designing assessment tasks. Others also used them to analyse tasks, ensuring that relevant specific features had been addressed and that sufficient marks had been allocated for each, which John reported as providing “a sense of balance” in assessment tasks. Adam suggested that their use to grade student work provided a checking on task validity. Harry reported that they contributed to designing “stronger tasks which are far more open-ended”. Peter believed performance standards had “raised the bar” for investigations and Russell believed their “key words” ensured tasks that “stretch the students at the top end”. These comments did not reflect what was observed in any of the quality management processes, nor could they be found in SACE Board documents. Instead, they appear to indicate participant reflection and evaluation of the new practices—transformative learning.

Between Rounds 1 and 3, the number of participants mentioning the advantage of performance standards in grading student work also increased, from zero to four. Elena, in Round 1, described them as helpful for “teachers who are inexperienced with marking”, which did not include her. In Round 3 she reported finding them useful for marking “subjective things”, such as practicals. Kevin reported faster marking of practical tasks when using performance standards. He used the time saved to give better feedback, which was beneficial to students. Brenda and Colin believed the performance standards provided evidence of the grades they awarded and used them in justifying grades to students.

Similar increased awareness of the advantages of performance standards occurred with moderators (from four in Round 1 to seven in Round 3) and non-moderators (from six to nine). This similarity and the diverse advantages mentioned suggested that appreciation arose from using the performance standards, rather than from involvement in the quality management processes.

Disadvantages of performance standards were addressed only in the interviews. Comparing the two rounds showed a decrease in mentions of disadvantages, from 32 in Round 1 to 24 in Round 3. The most common disadvantages cited were that performance standards are open to interpretation (ten participants in Round 1 and nine in Round 3) and are time-consuming (eight and seven mentions). The biggest

change was in criticism of their “subjective” nature, from six in Round 1 to two in Round 3. The wording of the performance standards was commonly criticised in both rounds, with them described as “airy fairy” and “wishy-washy”, but most commonly as “vague”. In Round 1 four participants saw a disadvantage in not being able to assign weights to “more important” features, falling to two in Round 3; this decrease may reflect that participants had achieved weighting by assessing some specific features more frequently. The criticisms of performance standards being time-consuming and open to interpretation did not decline over the year, suggesting that these had not been resolved by involvement in the quality management processes.

In summary, participants were able to cite purposes served by performance standards and advantages and disadvantages in their use. The common responses among participants suggested that increased knowledge regarding the purposes of performance standards might have arisen from involvement in the quality management processes. The diversity of advantages and disadvantage cited, however, suggested that awareness may have arisen from reflection on the use of performance standards—working with meaning (Moon, 2000)—or “reassessing the presuppositions on which ... beliefs are based” (Mezirov cited in Webster-Wright, 2009, p. 72)—transformative learning.

### ***7.2.2.3 Beliefs regarding social moderation***

Before 2011 Physics teachers submitted marks for SBA, which were statistically moderated against the examination; the introduction of social moderation was a major change. This section scrutinises the data for evidence, in the first and last interviews, of participant awareness of the advantages and disadvantages of social moderation.

Advantages were cited ten times in Round 1 and fourteen times in Round 3, with similar increases for both moderators and non-moderators. Four advantages were mentioned. Social moderation as good professional development changed from four mentions in Round 1 to five in Round 3. The invalidity of statistically moderating some tasks against an examination was mentioned twice in Round 1 and six times in Round 3. The perception that social moderation was fairer than statistical moderation changed from two to zero, while belief that social moderation forced

schools to undertake appropriate tasks changed from two in Round 1 to three in Round 3. Two participants knew teachers who had previously submitted “fictitious” marks for practicals and argued that this is not possible with social moderation. The observation about teachers’ practices was reinforced by an officer reporting one teacher asking others “How do you do a prac?”

Although the mention of advantages increased, the similar changes for moderators and non-moderators suggested no association with moderation experience, the only quality management process occurring between Rounds 1 and 3.

When asked about the disadvantages of social moderation, the number mentioned increased from 31 in Round 1 to 39 in Round 2. In each round six participants questioned the reliability of social moderation; for example, moderators changing their practices during the process and moderation outcomes differing between moderator pairs. The excessive time needed to undertake social moderation was mentioned by six in Round 1 and three in Round 3. Two disadvantages mentioned more than twice in one round were difficulty finding moderators (four in Round 1 and one in Round 3) and the unlikelihood of grades being raised (three in Round 1, two in Round 3). The unlikelihood of harsh markers having grades moderated upwards was mentioned by five different participants, with only one being a moderator. Thus there was no evidence that moderators perceived this as a flaw in moderation, although moderators’ opinions on this matter might not be considered unbiased. This matter is addressed in Section 8.2.1, which compares the SACE and Queensland systems. Some (one in Round 1, three in Round 3) expressed concern that social moderation divided teachers into two groups: those “in-the-know” (i.e., moderators) and others. Colin suggested that “collective knowledge and ... wisdom” would increase for moderators, while “those on the fringes [were] left out”. Four different participants (two in each round) commented on the “subjective” nature of social moderation.

The increase of eight in the number of disadvantages was isolated to moderators: ten in Round 1 to eighteen in Round 3. Increases greater than one occurred for the questionable reliability of social moderation, inconvenience to moderators and the lack of anonymity of moderated schools. It might be argued that moderation experience alerted participants to disadvantages of social moderation.



#### **7.2.2.4 Overview of changes in participants' assessment beliefs**

The data provided evidence of participants' beliefs regarding the summative use of SBA, the use of performance standards and social moderation. There was little evidence that involvement in the quality management processes might be associated with changes in beliefs regarding the use of SBA. The data suggested that knowledge of the purposes of performance standards might have arisen from involvement in the quality management processes.

Evidence suggested that some participants increased their awareness of the advantages and disadvantages of performance standards through reflecting on their own practices. The data also suggested that moderation experience might have contributed to awareness of disadvantages of social moderation and the possibility of SBA being open to manipulation.

#### **7.2.3 Familiarity with SACE Board terminology**

The current research sought evidence that teachers might have developed common understandings of concepts and terminology and that involvement in the SACE Board quality management processes might have contributed to such development. This section sought evidence of increased familiarity with SACE Board terminology over the research period.

A first approach scrutinised the transcripts for terms that had been re-defined for the new SACE: Kevin, for example, described the revision as partly “a change of jargon”. Redefined terms included *curriculum statements* becoming *subject outlines*, *learning outcomes* becoming *learning requirements*, *assessment components* becoming *assessment types* and *criteria for judging performance* becoming *assessment design criteria* (SACE Board of South Australia, 2010l, 2011d).

Identifying decreased use of a revised term would require its extensive use in the first round interviews. Round 1 transcripts showed no use of the terms *curriculum statement*, *learning outcome* or *outcome* (to describe student learning). The word *component* had been used four times. It was decided that none of these terms had been used sufficiently for comparison to Round 3. The term *criteria* (or *criterion*), however, had been used in Round 1 interviews by ten participants, rarely correctly referring to *assessment design criteria*. More frequently it had been used to refer to

*performance standards* or *specific features*. Adam, for example, said “one of the criteria talks about ‘consistently performing’, clearly referring to wording in the performance standards. On the other hand, Colin referred to the need to assess “the performance criteria ... [, for example,] “We need to assess KU1, we better put some KU1 questions in.” KU1 is one of the thirteen specific features, rather than one of the four assessment design criteria: Investigation, Analysis & Evaluation, Application and Knowledge & Understanding (SACE Board of South Australia, 2011d). Table 35 shows participant use of *criteria* in each round. The numbers show the frequency of use; lack of numbers indicate single participant use.

**Table 35: Participants’ meanings of *criteria* in interviews (n = 12)**

	<b>Round 1</b>	<b>Round 3</b>
Adam	performance standard	
Brenda		
Colin	specific features	
David		
Elena	specific features	
Frances	specific features	specific features performance standard
Greg	performance standard specific features	
Harry	specific features	specific features
John		
Kevin	performance standard	
Peter	specific features	assessment criteria
Russell	assessment design criteria (3) performance standard	assessment design criteria (3) specific features

Correct use of the term *criteria*, three times in each interview, was only made by Russell, who was not a moderator, and once by Peter, a moderator, in Round 3. Incorrect use was more widespread. In the first round *criteria* was used to refer to *performance standards* (four times) or *specific features* (six times). In the third round, incorrect use had fallen to once for *performance standards* and thrice for *specific features*.

The data might suggest that incorrect use had declined, from ten instances to four and from nine participants to three. Decrease in incorrect usage might suggest

association with greater involvement in the quality management processes, decreasing more among moderators (six down to two) than non-moderators (four down to two). Any conclusion, however, is dubious given the few participants and the few times the term was used.

A second approach analysed participant use of new terms: *specific features* and *performance standards*. Since 2011, Physics assessment has been based on four assessment design criteria (SACE Board of South Australia, 2011d, pp. 78, 79). These consisted of thirteen specific features and the performance standards describe five levels of student achievement for each specific feature. Table 36 gives two examples to illustrate the terms *specific features* and *performance standards* (SACE Board of South Australia, 2011d, pp. 84, 85).

**Table 36: Examples of specific features and performance standards**

<b>Specific feature</b>	I2: Selection and acknowledgment of information and data about physics and issues in physics from different sources	A1: Application of physics concepts and evidence from investigations to solve problems in new and familiar contexts.
<b>A grade</b>	Critically and logically selects and consistently and appropriately acknowledges information about physics and issues in physics from a range of sources.	Applies physics concepts and evidence from investigations to suggest solutions to complex problems in new and familiar contexts.
<b>B grade</b>	Logically selects and appropriately acknowledges information about physics and issues in physics from different sources.	Applies physics concepts and evidence from investigations to suggest solutions to problems in new and familiar contexts.
<b>C grade</b>	Selects with some focus, and mostly appropriately acknowledges, information about physics and issues in physics from different sources.	Applies physics concepts and evidence from investigations to suggest some solutions to basic problems in new or familiar contexts.
<b>D grade</b>	Selects and may partly acknowledge one or more sources of information about physics or an issue in physics.	Applies some evidence to describe some basic problems and identify one or more simple solutions, in familiar contexts.
<b>E grade</b>	Identifies a source of information about physics or an issue in physics.	Identifies a basic problem and attempts to identify a solution in a familiar context.

The transcripts of Rounds 1 and 3 were analysed for use of the two terms *specific features* and *performance standards*. The failure of participants to use the term *specific feature* was noteworthy, being used correctly once by each of two participants in Round 1 and twice correctly, by a different participant, in Round 3. Kevin, who had used the term correctly in Round 1, in Round 3 asked “what are the little bits called?”

Scrutiny of the transcripts showed that the term *performance standards* was used in three ways. First, it was used by all participants to refer to the new assessment system. This was also common among SACE Board officers and was not investigated further. Second, it was used, as in SACE Board documents, to describe

student achievement. Third, it was used to refer to the specific features rather than student achievement in those features. Table 37 shows the number of times each participant correctly used the term *performance standard* to refer to student achievement and incorrectly referred to *specific features*.

**Table 37: Participants' use of the term *performance standard* (n = 12)**

	Referring to student achievement (correct)		Referring to specific feature (incorrect)	
	Round 1	Round 3	Round 1	Round 3
Adam			2	
Brenda	1	2	9	6
Colin	1	2	1	
David	7	7	6	4
Elena		1	2	
Frances	5	5	6	4
Greg		1	1	11
Harry		7	3	2
John	6	11	3	1
Kevin	5	4	1	3
Peter	3	3	1	2
Russell	1	1		1
<b>TOTALS</b>	<b>29</b>	<b>44</b>	<b>35</b>	<b>34</b>

Correct use had increased (from eight participants to eleven and from 29 occurrences to 44), while incorrect use had declined (from eleven participants to nine and from 35 occurrences to 34). On examination, the six participants (Adam, Brenda, David, Frances, Harry and John) with the greatest increases in correct usage, or the greatest decreases in incorrect usage, included three moderators and three non-moderators. It was reasoned, therefore, that changes in usage of the term *performance standard* did not appear to be related to involvement in the quality management processes, but rather to teachers' personal characteristics beyond the scope of the present study. It is argued (see 8.4) that improved familiarity with terminology requires ongoing usage by all teachers.

#### **7.2.4 Consistency in application of performance standards**

The development of shared understating of standards has been reported as an outcome of social moderation (Adie et al., 2012; Connolly et al., 2012; Hipkins & Robertson, 2011; L. Reid, 2007). Although the current research was not restricted to social moderation, data were scrutinised for evidence of improved consistency in the application of the performance standards. Opportunities for improving consistency occurred in the clarifying forums and moderation.

After the clarifying forums, participants were asked whether they agreed or disagreed that the forums had helped in applying the performance standards in a manner consistent with other teachers. Four disagreed, one strongly; four were neutral and five agreed. Overall, it appeared that participants did not acknowledge that clarifying forums had generated consistency among teachers in the use of performance standards.

After moderation, moderators were asked the same question. One strongly agreed, four agreed and one was neutral. Five reported that disagreement between moderators was rare and generally restricted to differences of one on the SACE Board 15-point scale. Adam believed that changing partners each day had helped to develop a common standard, although two believed that more training would have improved consistency. SACE Board officers reported consistency among moderators, based on benchmarking results and observations and reports of moderator discussions. It appeared that consistency existed among moderators but that it might have been improved.

In the verbalisation activity, participants were asked to grade the work of two students in Rounds 1 and 3. Data were restricted to the four moderators and four non-moderators who completed both activities.

Round 1 occurred before moderation and Round 3 after moderation. Table 38 shows the results of the verbalisation exercise.

**Table 38: Grades for verbalisation exercise**

	Student 1		Student 2	
	Round 1 Grade	Round 3 Grade	Round 1 Grade	Round 3 Grade
<b>Moderators</b>				
Harry	B+	B-	A-	A
Kevin	C-	B-	B-	B+
Peter	B	B-	B	B+
Greg	C	C	C	B+
<b>Range</b>	5	2	6	2
<b>Non-moderators</b>				
Brenda	D	C	B-	B
Colin	B-	B-	B+	A-
John	C+	D+	B-	B
Russell	C-	C+	B+	B
<b>Range</b>	5	4	2	2

The grades assigned used the SACE Board 15-point scale. The ranges of moderators' grades narrowed from 5 to 2 (Student 1) and 6 to 2 (Student 2), suggesting moderation participation had improved consistency. For non-moderators the range for Student 1 narrowed from 5 to 4, less improvement than for moderators. For both groups, this improvement might have been a consequence of the year of working with performance standards in teaching and assessment. Non-moderators' grades for Student 2, however, confound the picture. Their grades were consistent (within a range of 2) in the first round and this consistency was maintained. This was equivalent to the consistency shown by moderators after moderation.

It was considered that, while some data might suggest improvement, overall, they were insufficient to support claims that involvement in social moderation had improved grading consistency. The small number of participants weakened any inference.

### 7.2.5 Teacher assessment confidence

Since increased teacher confidence has been reported as an outcome of teacher learning (Limbrick & Knight, 2005; L. Reid, 2007), the current research collected data on teacher confidence. Data from the first and final interviews are considered

below for six moderators and six non-moderators who completed both interviews.

Among the non-moderators, Brenda had found 2011 worrying, possibly because she worked alone. She constantly wondered: “Am I doing it right?” Using the 85–70–55 cut-offs reassured her that grades were unlikely to go down as cut-offs were higher than with the 15-point scale. She worried, however, that she may have graded students too harshly. When her grades were confirmed she was reassured but “not 100% confident yet”, wondering whether moderators had “only skimmed through” the work submitted. Colin had been “reasonably confident” that his grades would be confirmed and was puzzled when two classes had different outcomes, even though cross-marking on the same tasks had been carried out. Consequently, he questioned whether he had caused the change or whether it was due to “some aberration in the system”. The limited nature of feedback was not helpful and he was now less confident than before moderation. Frances was also puzzled at the different outcomes for the two classes in her school, given their common tasks and cross-marking. With extensive SACE Board experience, she was confident in her own assessment. Her experiences in the development of the new SACE, however, convinced her it was “a substandard system”. Consequently, she attributed the change to moderators “stuffing up” their results. Three non-moderators were reassured when their grades were confirmed at moderation, while one’s assessment confidence was unchanged.

Of the six moderators, Adam was “pretty confident” prior to moderation, but slightly concerned about possibly jeopardising students’ grades. Kevin had been “extremely confident” that he had been consistent with previous years. He was less confident about “what moderation may do”, as were two others who were confident about their own assessments. All four were reassured or “relieved” when moderation confirmed their grades. Before moderation, Greg, because of his examination marking experience, had more confidence in his tests than his investigations. He was not confident in using the performance standards in grading. At moderation he was concerned at differences between moderators but was consoled that, as a moderator, he was better informed than non-moderators. All moderators described similar self-confidence arising from the belief that they were “in the know”. They were somewhat reassured by knowing they had not made mistakes evident in other



teachers' task designs and gradings. Moderation experience also bolstered self-confidence by seeing good practices that could be adopted.

For both moderators and non-moderators, it was confirmation of their grades that generated assessment self-confidence. Where grades were not confirmed, the lack of informative feedback contributed to a decline in self-confidence. All participants demonstrated what could be considered a passive position: moderation outcomes affected how they perceived themselves, their "professional identity" (van den Berg, 2002, p. 582).

### **7.2.6 Teacher attitudes**

The first and last interviews were scrutinised for positive and negative comments to find whether there had been any change in attitudes over the research period.

Because twelve participants agreed to both interviews, their transcripts were used in the analysis.

In Round 1 interviews, eleven participants made a total of 35 negative comments, whereas in the final interviews a total of seven negative comments were made by six participants: decreases in the number of participants offering negative comments and in the number of negative comments. The targets of the comments also changed. In the first round, uncertainty was paramount (mentioned by six participants), followed by work load (four participants), frustration with the lack of information (three participants) and concerns about moderation (three participants). At the end of the research period, frustration at poor moderation feedback was the major concern (five participants).

While negative comments decreased over the research period, positive comments increased from ten (made by seven participants) to 21 (made by nine). In the first round, most of the positive comments (made by four participants) reflected outlook. Some, for example, mentioned that they welcomed change, or were prepared to make the best of it. In the third round, moderation outcomes were the major factor stimulating positive comments. For six participants, the minimal (or absence of) changes to grades generated positive attitudes, while moderation experience gave confidence to three.

It appeared that participants had become more positive towards the changes in the assessment system, with moderation involvement and confirmation of grades in moderation being important in generating more positive attitudes.

### **7.2.7 Overview of evidence of participant change**

The evidence suggested that participants had changed over the research period. Without a rationale for the adoption of performance standards, participants had, however, used them, more in assessment task design than in grading student work. Changes in assessment confidence appeared to be externally derived, linked to moderation outcomes rather than experiences during the year. Experience—learning through practice—appeared to be the greatest contributor to participant learning, raising awareness of the benefits of performance standards and of weaknesses in social moderation, as well as improving usage of new terms. Any potential for teacher learning inherent in quality management processes might be better realised if the process were modified to integrate them more closely into teachers' daily work. Recommendations are made later (see 8.4).

## **7.3 The nature of teacher learning**

The current research examined the data for evidence of the nature of participant learning. Approaches to learning are discussed in terms of their surface, strategic or deep nature (Entwistle, 1998; Moon, 2000). A *strategic* approach, “alertness to assessment requirements” (Entwistle, 2000, p. 2), was observed in participants ensuring compliance with the perceived expectations of moderators. A *surface* approach referred to participants undertaking what was expected of them, without questioning or attempting understanding (Entwistle, 2000; Marton & Saljo, 2005; Moon, 2000). A *deep* approach was demonstrated by participants looking for principles and relationships between ideas and in questioning the assessment system and their own practices and beliefs. Learning of a deeper nature is discussed with reference to Moon's (2000) stages of learning—making meaning, working with meaning and transformative learning. Data were also scrutinised according to whether participants assumed active or passive roles. When adopting an active role, learners “manifest a desire to find things out, ... asking questions ... [and taking] ownership of their learning” (Crick, Broadfoot, & Claxton, 2004, p. 255); they reflect critically on their experiences (Webster-Wright, 2009). In contrast, a passive role

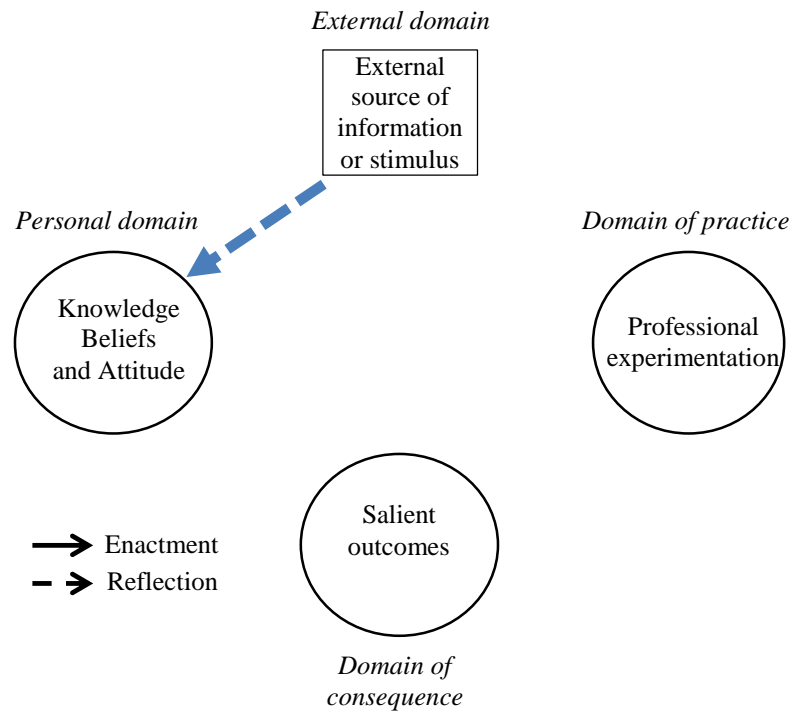
would be indicated by failure to ask questions or take initiative and by uncritical acceptance of information provided (Crick et al., 2004); the difference lies in whether initiative lies with the teacher—an active role—or with some other agency—a passive role. The data were also scrutinised for evidence of change sequences (D. J. Clarke & Hollingsworth, 2002) that described participant learning.

Because each learning journey was unique, a summative description was not considered appropriate. Equally, to describe all would involve considerable repetition. Consequently, it was decided that the journeys of three would be described, with references to others' where appropriate. Similarly, to avoid repetition, the interconnected model of teacher learning (D. J. Clarke & Hollingsworth, 2002) has been used to illustrate a selection of the change sequences in teachers' practices, attitudes and beliefs. Nevertheless, the model has been used to draw some general conclusions regarding teacher learning.

### **7.3.1 Brenda's learning journey**

Brenda had been teaching for twenty years, with a decade of experience in senior Physics. She had no previous involvement in SACE assessment panels and this continued during 2011. She was the only Physics teacher in her school and contact with other Physics teachers was limited to social meetings during school holidays. She was very conscious of working alone and, although others suggested contacting the SACE Board, she claimed "I don't know enough even to know what to ask" and suspected the Board's response would be to "talk to other Physics teachers". Consequently, believing "it's hard enough to talk to people in the school, let alone to find somebody else", she did not contact the SACE Board, assuming a passive role.

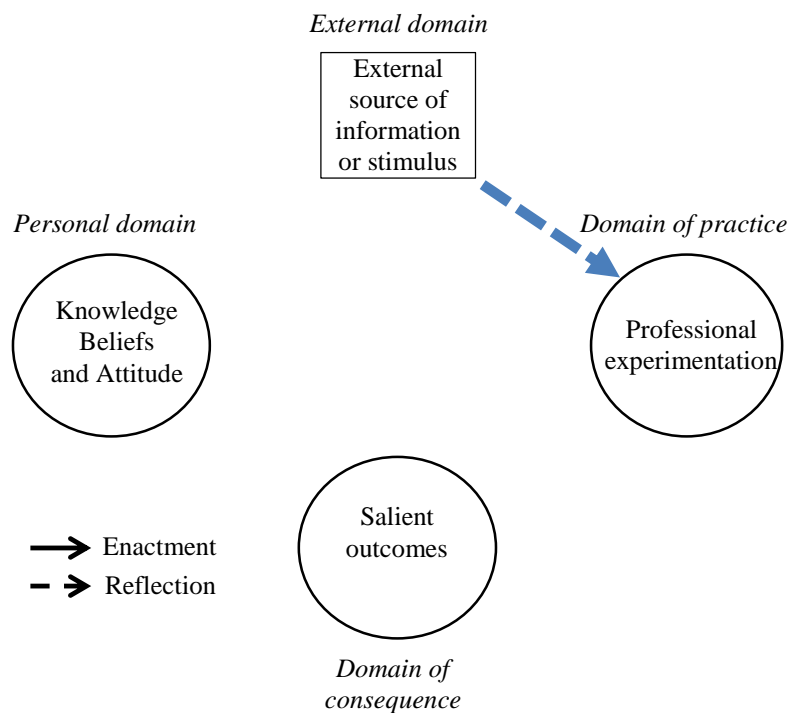
She was, however, active in taking her tasks to the clarifying forum and showing them to other teachers. Comments "that looks OK" provided some assurance that her practice was acceptable. Figure 10 shows this action as a change sequence using the interconnected model of teacher change (D. J. Clarke & Hollingsworth, 2002).



**Figure 10: Change sequence Brenda—consulting other teachers**

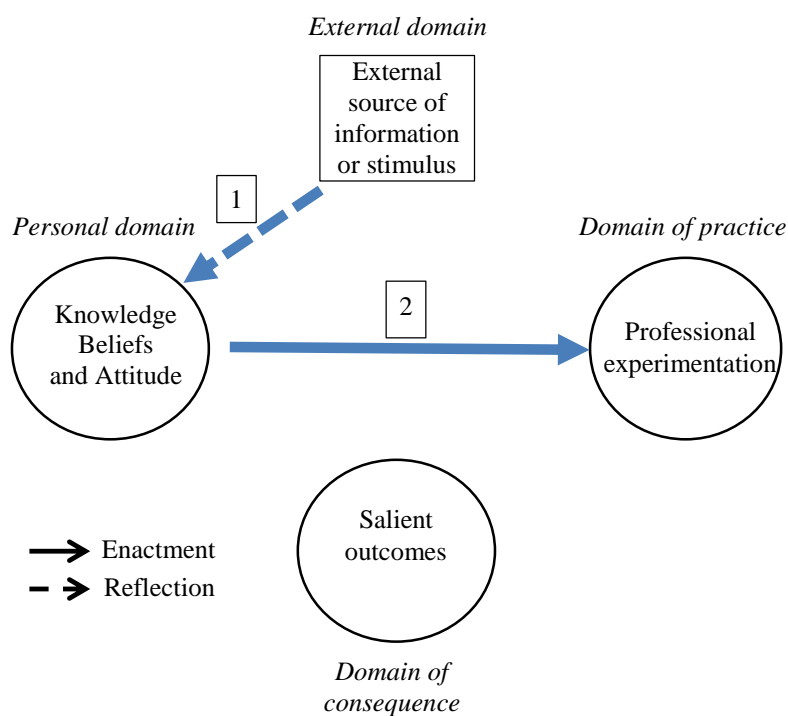
In this figure the external stimulus is the teacher response to Brenda’s tasks. Reflection on this, shown by the arrow, generated a change in belief: increased confidence in her assessment.

“Inventors often build their own wheels by using other people’s spokes” (Popham, 1988, p. 22) and this was evident in the current research, with participants making use of external sources. Figure 11 shows an example of a change sequence in which a teacher enacts a practice from an external source.



**Figure 11: Change sequence—enactment of externally modelled practice**

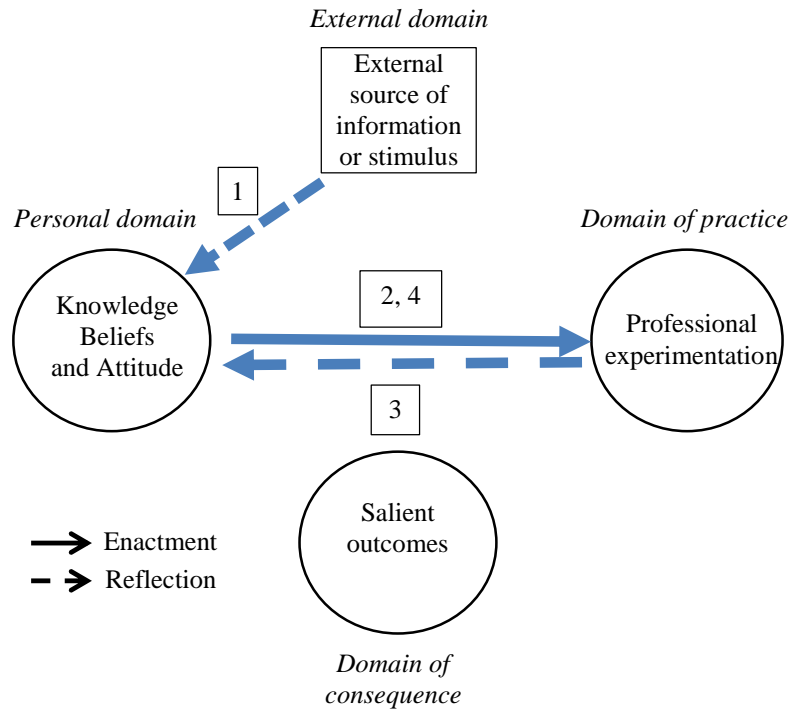
Brenda approached this enactment by “virtually plagiarising” the Issues Investigation exemplar from the SACE Board website. Even so, by adapting it “to make sense in my head”, she was doing more than Moon (2000) ascribes to this same term. Rather than examining the exemplar “in relation to itself” (Moon, 2000, p. 142), she was relating the exemplar to her understanding of the performance standards, making meaning and using her understanding “towards a particular purpose” (Moon, 2000, p. 139), working with meaning. Figure 12 shows this change sequence.



**Figure 12: Change sequence Brenda—exemplar use in task design**

Arrow 1 shows the reflection on the website exemplar (the external source) which provided information in the personal domain. The enactment (Arrow 2) shows Brenda designing her task, in which she assigned specific features to each component.

Having used the task in 2011, she considered students' responses, reflecting on its success (Arrow 3) and revised it for 2012 (Arrow 4) as shown in Figure 13.



**Figure 13: Change sequence Brenda—exemplar use in task design**

There was only one other instance of a participant using an exemplar with minimal adaptation. Elena had used the LAP exemplar from the SACE Board website and her experience can also be illustrated with reference to Figure 13. She used the exemplar to increase her knowledge (Arrow 1) and, changing tasks slightly because of equipment availability, used the website exemplar to write her LAP (Arrow 2). When designing tasks, she found that some of the specific features shown in the LAP were inappropriate (Arrow 3) and revised the LAP for 2012 (Arrow 4).

Much of Brenda's learning during 2011 fitted the change sequence of Figure 13, sometimes with multiple reflections (equivalent to Arrows 2 and 4) and enactments (equivalent to Arrow 3). For example, reflection on the new syllabus (stimulus from the external domain) initiated a belief that specific features needed to be used in designing tests (Arrow 1). She adapted previous tests by assigning specific features to each question (Arrow 2) with, for example, a question worth 3 marks being counted as 3 marks for multiple specific features; thus a test out of 90 marks might be out of 120 after multiple assigning. Upon reflection (Arrow 3), Brenda limited any question to a maximum of two specific features (Arrow 4). Later, in another reflection-enactment sequence (Arrows 3 and 4), she assigned marks to specific features so that they added up to the marks assigned; thus a question worth 3 marks

might have 2 marks assigned to one specific feature and 1 mark to another.

All participants assigned specific features to questions in various ways, with many subsequently changing their procedures. Some participants used this procedure to check the test, whereas others assigned specific features during test construction. There were also differences in the number of specific features assigned to questions: Frances and John, for example, assigned only one specific feature to a question. Some who had assigned multiple specific features reduced the number later, as had Brenda, with one declaring that doing so “makes life a little bit easier”.

When Brenda graded tasks, she mathematically determined the grade for each specific feature, reasoning “otherwise it’s all too subjective”. Showing students the percentage for a specific feature served as “objective evidence” for the grade assigned. At no stage in her grading was reference made to the wording of the performance standards. The wording was only used prior to submitting materials for moderation when she reviewed grades for marks just below her cut-offs and adjusted some upwards (e.g., changed a D+ to a C-). Her faith in numerical grade determination reflects literature reports that mathematics and science teachers are uncomfortable using standards considered to be “open to interpretation” (Klenowski & Wyatt-Smith, 2010a, p. 34) and are more comfortable with marks. This appeared to be supported in the current research: all participants preferred numerical procedures for grade determination.

At the end of the research period, Brenda reflected on the changes to the SACE and acknowledged they had “made me think about what I’m giving the kids” (i.e., task design). She believed that addressing the performance standards had improved her tests: more explanatory answers were required whereas previously the focus had been on calculations.

Brenda blamed her “maths soul” for being irritated by the specific features not being evenly weighted. Although believing that the use of performance standards implied equal weighting, she found that there was “no way in this world” she could achieve equality. She ensured, however, that each specific feature had at least 10 marks in each test out of 90, thus achieving some semblance of equality. Weighting of specific features was an issue for all participants, who resolved the issue differently.



Some, like Brenda, attempted to make them equal whereas Russell, for example, made no attempt, arguing “they’re not, in my eyes, equally important”. He achieved *de facto* weighting by assessing some specific features more times than others. Thus, when he holistically scanned highlighted performance standards in generating grades, features highlighted more frequently contributed more.

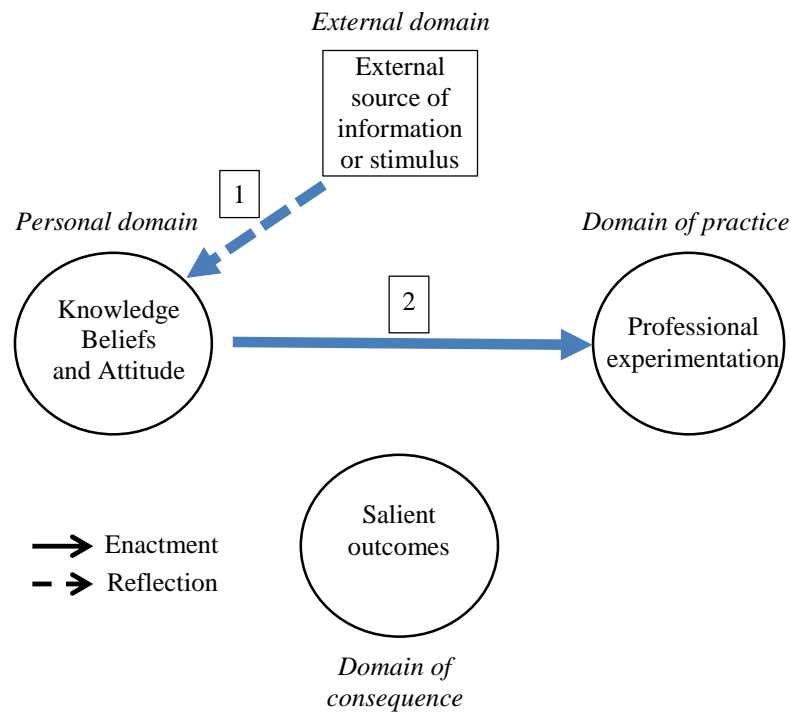
Brenda contrasted her working alone to previous schools where she had frequent discussions with colleagues. Holiday discussions with others were less fruitful; colleagues were “desperate”, more confused than she was. The fact that only one student had a minimal grade change at moderation provided some confidence that she was “on the right track”.

In reflecting on the year, Brenda did not believe she had at any stage evaluated whether the changes represented an improvement, explaining: “this is what I have to do and I’ll just do it”. She did, however, believe her tasks were better and acknowledged “probably we should have been doing it before but we thought we were”. Her use of a spreadsheet and “having to circle rubrics” required “massive time” and she described the outcome as “taking substantially longer for slightly better”. It might have been that, with no opportunity for regular sharing, she was hesitant to acknowledge her achievements.

### **7.3.2 John’s learning journey**

John had been teaching for over thirty years, all involving senior Physics. He described, as did all participants, an evolution of his assessment practices and understandings. He worked closely with the other Physics teacher in the school and used the words “we and “our” frequently in interviews, for example describing how “we find better ways of doing things and ... part of it is throwing away our old emphases”. As reported previously, he collaborated with a group of teachers in nearby schools.

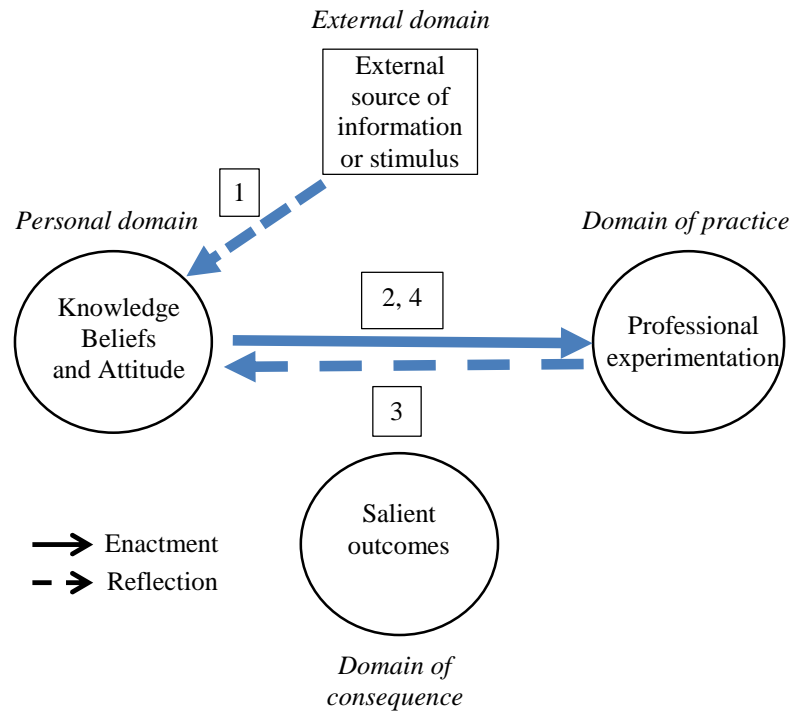
The clarifying forum stimulated learning for John, as shown in Figure 14 using the interconnected model of teacher learning (D. J. Clarke & Hollingsworth, 2002).



**Figure 14: Change sequence John—the importance of task design**

In the forum (in the external domain), an apparently perfect practical report had been awarded a B grade. Reflection on this (Arrow 1) generated belief in the “importance of constructing your tasks ... [to give] top students ... a chance to fly”. This was enacted by modifying practical tasks, for example, by ensuring that students had opportunities to make predictions and by increasing the analysis requirement. All participants, at some stage, acknowledged the importance of task design as something that had been learned from the SACE changes. Elena explained that the students needed to “introduce the complexities” in a report and was careful not to provide too much information. Greg and Frances grappled with the need to inform students about the assessment requirements without making tasks “too scaffolded”.

John explained how the performance standards had allowed identification of students’ strengths and weaknesses. For example, “the knowledge and understanding for some students was really quite high but the analysis ... was not”. This knowledge informed reports written for the students. Figure 15 shows this change sequence.



**Figure 15: Change sequence John—student information from performance standards and using performance standards to write reports**

In Figure 15 the change in the external domain was the information that performance standards were to be used for providing student feedback (SACE Board of South Australia, 2011d). Arrow 1 shows reflection generating a belief that performance standards might provide students information about their strengths and weaknesses. Arrow 2 shows the enactment, in which performance standards were used with student work. Arrow 3 shows the reflection that using performance standards had generated useful information and Arrow 4 shows the change in practice—the use of performance standards to generate feedback to students. Although acknowledging that the performance standards improved feedback, John argued that an “enormous amount of time” had been spent that could have been better spent on supporting student learning. Hence there was no change in the domain of consequence: John did not value the feedback as a “better way” (Evans, 2014). Evans (2014) argued that imposed behavioural changes did not require recognition of a “better way” and did not constitute teacher learning. John likened the professional development associated with the new SACE to training doctors “to operate with a new system of Medicare”, with no improvement in patient care: the teacher professional development associated with the new SACE had not “enabled us to ... engage with the students and encourage their learning any better”. John had acknowledged the

importance of designing tasks that stretched students (see Figure 14), but did not recognise this as encouraging better learning.

John attempted using the performance standards when writing reports for parents and students, a change sequence identical to that shown in Figure 15, and common among participants. In this instance, the external domain change was the syllabus advice. Arrow 1 shows reflection that generated a belief that performance standards might be used to write reports for parents and students. Arrow 2 shows the enactment: report writing using the performance standards. Reflecting on this (Arrow 3), John considered the reports to be “pretty indigestible prose”, leading to a revised belief about using performance standards in writing reports. Arrow 4 shows the enactment of the new belief: performance standards, although useful in ascertaining students’ strengths and weaknesses, needed to be rephrased for reports. Although most participants acknowledged the usefulness of the performance standards in defining skills expected of students, criticism of their wording was common.

John’s changes in the marking of practical work provides another example of evolving teacher practice. The earliest were marked and the marks converted to grades. They then assigned specific features to “each of the elements” (e.g., aim, results and conclusion), marked each specific feature and assigned grades based on the marks. He described this as “a hugely difficult exercise” and consequently, towards the end of 2011, they assigned grades for the feature without using marks. This was revised in 2012 when both teachers reverted to assigning marks for specific features, believing this to be more precise than the holistic assigning of grades. From the first interview, John had been concerned that performance standards “don’t allow the fine grained assessment ... appropriate” for tertiary entrance purposes. It was this belief that underlay the return to the spreadsheet, which “makes us feel comfortable that we’re doing it correctly”.

John doubted whether the use of performance standards and specific features had generated better-constructed tests, arguing that tests were constructed as previously. He did, however, acknowledge that practical tasks had been “extended ... to ensure that the top students had the ... [opportunity] to show their ability”.

John had done “a fair bit of reading” on the use of performance standards and had misgivings around their “blunt nature”. He argued that “by their very nature ... performance standards produce ordinal data” and was critical of the SACE Board combining SBA ordinal data with the continuous data from the examination. Kevin also argued that the use of ordinal data to generate ATARs was “mathematically flawed”, arguing that Physics students would be penalised for using a measurement “as vague as a performance standard ... [to generate] a number with four significant figures”, that is ATARs. John would have preferred a 100-point scale for the SBA, providing “more fine-grained” data. He was “more comfortable” marking parts of tasks rather than grading holistically, similar to Kevin’s preference to make “a whole lot of small judgements”.

In ‘making meaning’, new material “is related to that which is already known” (Moon, 2000, p. 139). At the end of the year, John and colleague made a conscious decision to use previous years’ results to inform the grades they submitted to moderation. They mentally compared students with past students and, “guided by our experience”, graded their current students. John appeared to be the only participant who made such a conscious comparison to previous students. Greg had used marks to rank his students and determined grade cut-offs “from experience”. The other nine with previous SACE experience appeared to have made more intuitive comparisons with previous years, with mixed responses. Two believed that standards for SBA were similar to previous years, four believed that it was now easier to gain a C grade but that an A grade was more difficult to achieve. Three believed the expectations had been raised for investigations but were unchanged for tests. All had, however, undertaken some relative comparison with the previous system—transformative thinking—but used different levels of complexity in their analyses.

When asked about important events over the year, nothing particular stood out for John: his acknowledged affordances included intra-school and inter-school collaborations and attending SACE Board meetings. While much intra-school collaboration was with the other Physics teacher, he described discussions with other teachers and used the term “triangulation” for his “looking at the same thing from different perspectives”. The only constraint he suggested was the belief within the

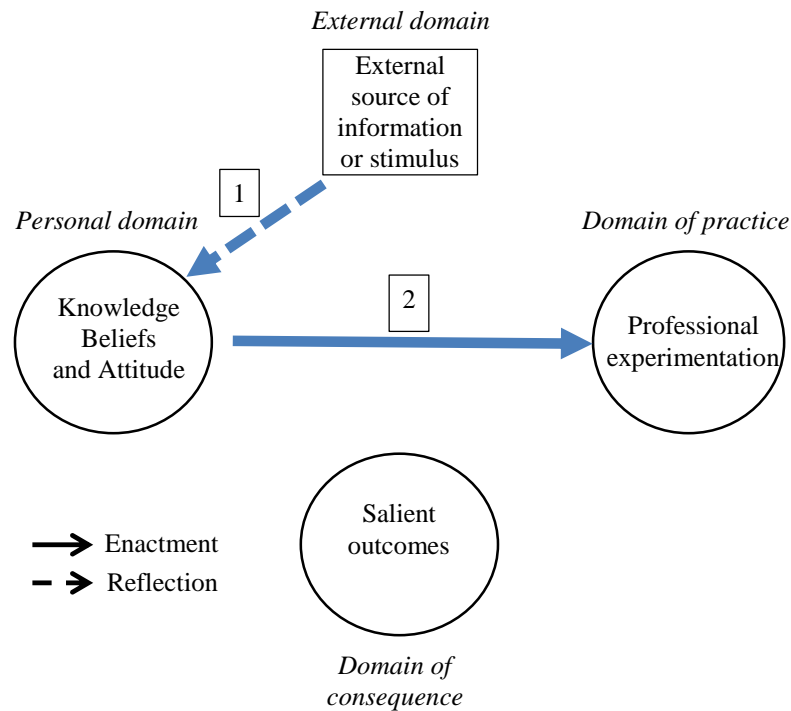
school, “which is focussed on university entrance”, that the changes in the SACE were due to concern about SACE completion rates not relevant to his school.

John had used specific features in designing assessment tasks and had used performance standards in grading student work and for informing reports. At the end of the research period, however, he was still concerned at their imprecise nature. He explained: “The performance standards have helped us to change the way we mark. I’m not sure that they have made it any easier or difficult or better. It’s just different.” John stated that the performance standards had “forced a rethink on assessment” but believed that “on reflection the experience has affirmed most of my previous practice”. It would appear that John had reflected greatly on the changes in the SACE but had not changed his beliefs, although it might be suggested that he had “tweaked” them. It might be argued that his interrogation, and subsequent affirmation, of his own beliefs and practices may not have occurred without the SACE changes.

### **7.3.3 Kevin’s learning journey**

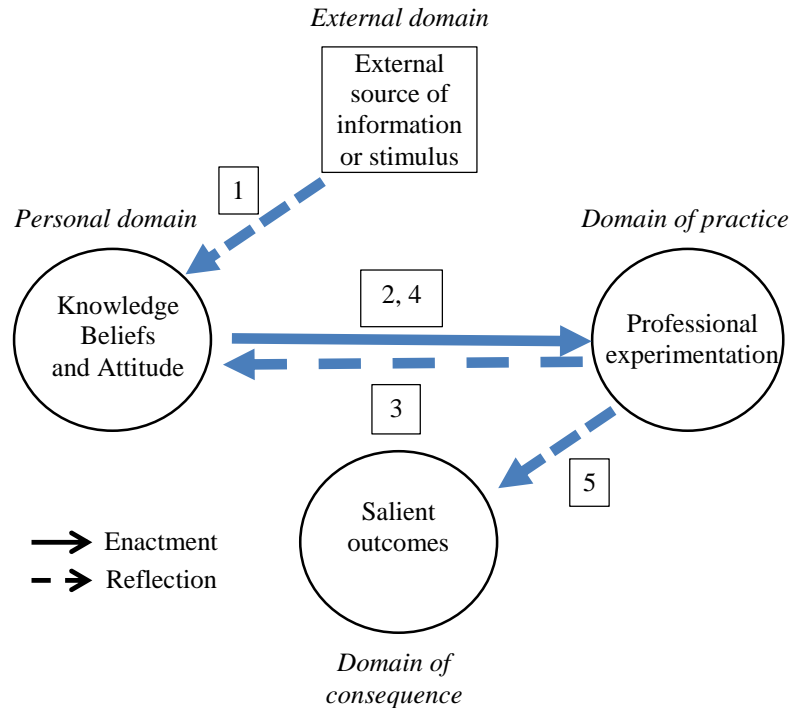
Kevin had been teaching for over twenty years, with more than ten years’ experience of Senior Physics teaching, SACE marking and SACE moderation. His pre-existing collaboration with his other Physics teacher continued and he was a member of the most effective inter-school collaboration encountered in the current research.

Kevin explained that his treatment of practical tasks had changed most because he was no longer “thinking marks”. This had evolved over the year as shown in Figure 16.



**Figure 16: Change sequence Kevin—assessment of practical work Phase 1**

Early in 2011 the two Physics teachers had reflected (Arrow 1) on the external stimulus, the revised syllabus, and believed that the specific features and performance standards needed to be addressed in practical work. This was enacted (Arrow 2) by assigning specific features to the components (e.g., aim, hypothesis, results) of previous tasks. Some specific features (e.g., recording data) were assigned more marks than others (e.g., collaboration). This was then modified as shown in Figure 17.



**Figure 17: Change sequence Kevin—assessment of practical work Phase 2**

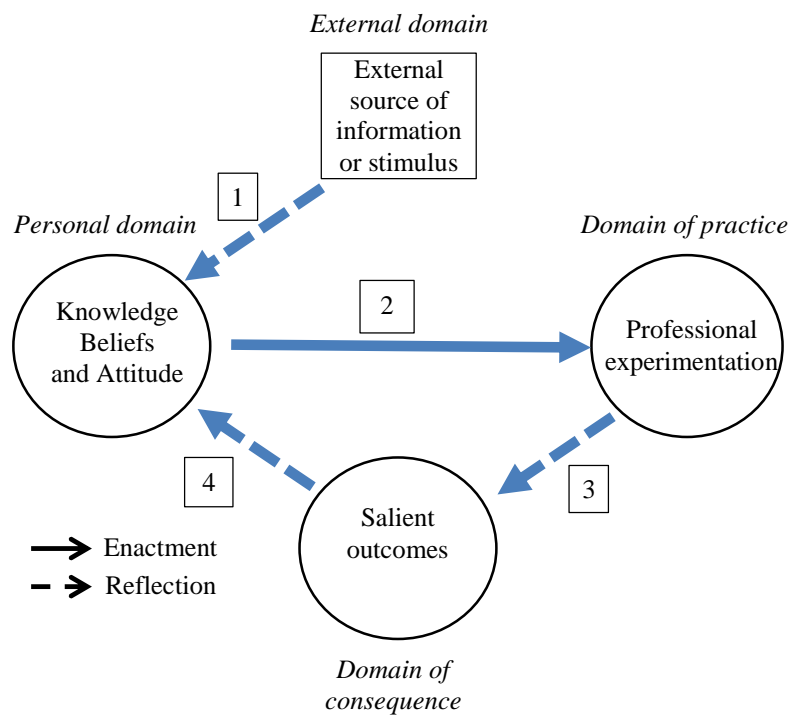
The inter-school group discussed the assessment of manipulative and collaborative skills (external source). Reflecting on the discussion (Arrow 1), Kevin and his colleague believed they needed some record of their observations of student skills. They enacted this (Arrow 2) by including items on their task sheets: a table listing specific features with marks and adjectives (e.g., *poor*, *acceptable*, *skilful*) and instructions for students to include a self-report on their collaborative practices. They used the tables to assign marks out of five for each specific feature. Reflecting on this practice (Arrow 3), they removed the “superfluous” marks and assigned grades for specific features based on the wording of the performance standards (Arrow 4). Further reflection (Arrow 5) recognised faster marking as a benefit of performance standards: “You spend more time designing the task but then ... you’re making a holistic judgement ... rather than adding up two marks and three marks and one mark, etc.” By the end of the research, eight participants used the wording of the performance standards to grade investigations, describing them as “quite useful when marking subjective things”. Although Kevin acknowledged the marking speed afforded by performance standards, he did not acknowledge them as an improvement on the previous system without performance standards: they were still described as “broad and vague, ... not designed for fine judgement”. The argument that



behavioural change, without recognition of a “better way”, does not constitute teacher learning (Evans, 2014), is discussed in section 8.5.5.

In another sequence of reflection and enactment (a repeat of Arrows 3 and 4) Kevin and his co-teacher found it necessary to make additions because some specific features (e.g., collaboration) had not been previously addressed. Other participants also mentioned how performance standards required assessment of skills that might previously have been given a lower priority: experimental design, error analysis and safety. Harry believed that the performance standards had raised the standards of practical tasks: reports now incorporated much more analysis. Peter began teaching students how to write a practical report.

Like Adam (see 3.9.5.1), Kevin changed his mind about LAPs, with the change sequence shown in Figure 18.



**Figure 18: Change sequence Kevin—belief in usefulness of LAPs**

Initially Kevin believed that the LAP served only a bureaucratic purpose, whereas in the final interview he acknowledged a change of mind. The external domain stimulus was the requirement to have the LAP approved by the SACE Board. Reflection (Arrow 1) led him to believe that its completion was merely bureaucratic.

Enactment (Arrow 2) required the development of a LAP that was submitted and approved. He realised, however, that the plan helped ensure that all syllabus requirements were addressed: “it’s a way of making sure the teacher keeps that in mind as they’re designing their tasks”. Thus he valued the outcome (Arrow 3) and reflection (Arrow 4) led to his changed belief: the LAP was “more than bureaucratic”.

Kevin reported a number of instances in which he evaluated the changes in the SACE and in his own assessment practices. He saw benefit in the tables mentioned earlier for recording assessments of practical skills, but was dismissive of instructing students to write self-reports on their collaborative skills. He acknowledged these as solely for moderators, describing them as “just a nonsense”: students were being assessed on their “ability to explain themselves”, rather than the ability to work collaboratively.

Kevin described moderation as “different learning” from examination marking: a shift from teaching students “how to answer questions” for examinations, to teachers learning “how to ask the right questions”. Thus moderation was a valuable learning experience in task design.

He reported that the use of performance standards had “filtered down” to earlier years in the school for two reasons: they were faster to mark and pre-SACE students became familiar with them. Other participants also reported that their use in SACE had influenced practices in the middle school. Harry described how “more rigorous” practical work was being introduced in Year 8 and that scaffolding was being introduced to develop students’ report-writing skills, while Colin reported likewise for the collaborative skills. Thus, although participants had been reluctant in their acceptance of performance standards, a number voluntarily began using them with younger students.

Kevin described moderation as “big” in his learning for 2011. He had spent the year “stressed” that he might have done something that jeopardised students’ results. Moderation increased his confidence; he had seen “mistakes that other teachers were making” that he had not made and moderation had confirmed his grades.

Reflecting on the year, Kevin was very conscious of the benefit of working with

others: “To do all this on my own would be so scary.” The most important affordance was working closely with the other Physics teacher. There had been some sharing with others in the school, including a meeting where Science teachers shared assessment tasks. Without further meetings, however, “the sciences have actually grown apart”. He also acknowledged the value of inter-school collaboration, SACE Board meetings and exemplars, as well as officers’ responses to questions.

The only constraint identified by Kevin was previous experience of the new SACE, in which he was involved. He was “pretty peeved” that the final document bore “very little resemblance” to that produced by the teachers. Seeing subject teachers’ input ignored, he had “got out of all the SACE committees” arguing “if they’re not going to listen to me ... there’s no point me doing it”.

#### **7.3.4 Participants’ reflections on learning**

In the final questionnaires and interviews, participants were asked to reflect on learning that might have occurred during 2011. The questionnaire included items aimed at identifying steps in teacher learning. Table 39 shows the frequency of responses.

**Table 39: Frequency of steps in learning**

<b>Statement</b>	<b>Frequency (n = 13)</b>
I have questioned the way I carry out assessment	12
I gathered information I needed to adopt new ways of assessing	11
I tried to figure out a way to adopt new ways of assessing	10
I realised I still agreed with my original beliefs	9
I thought about carrying out my assessment in a different way	9
I tried out new forms of assessment so that I would become more comfortable or confident with them	9
I took action and adopted new ways of assessing	9
I changed my ways of assessing a number of times	8
I have questioned my beliefs about assessment	8
I tried out new forms of assessment so that I could select the best format	8
I realised that other teachers were also questioning their beliefs about assessment	7
I began to think about reactions and feedback about my new forms of assessment	6
I felt uncomfortable with my previous form of assessment	1
I realised I no longer agreed with my previous beliefs	0

While the items with the highest frequency present a picture of teachers grappling with change, responses to the final two items suggest that participants may have retained their original assessment beliefs. This was supported in interviews. When asked whether they had changed their beliefs about assessment, most participants stated (sometimes quite firmly) that they had not. Previous research has shown that teachers do not easily identify significant changes in beliefs, even when able to identify changes in practice (Bissaker, 2009).

By the end of the research period, eight participants were using the wording of the performance standards to grade investigation tasks. They had changed from numerical marking when not obliged to do so. In responding to the last statement in Table 39, however, they claimed that their beliefs were unchanged. It appeared that most of the participants had changed their practices in what appeared to be important ways, although disinclined to admit to a change in beliefs. It might have been that, as teachers of a subject at the “‘high-end’ of the curriculum” (Crafter et al., 2006, p. 41), assessment was “so closely tied with the rewards and penalties of their jobs, they

could not afford to let loose of their original beliefs” (Yerrick, Parke, & Nugent, 1997, p. 155). Alternatively, it might suggest that they were in a state of change but not ready to acknowledge it, appearing to support studies reporting that, faced with “externally imposed expectations”, teachers’ responses can be seen as “ambiguous ... and even contradictory” (van den Berg, 2002, p. 580). Any change in beliefs in the current research would support Guskey’s (1986) model of teacher change—staff development caused changes in teacher practices prior to changes in their beliefs and attitudes—rather than the “conventional” model in which change in beliefs precedes change in practice (Loucks-Horsley et al., 2003).

In the final interview participants were asked to comment on important events during the year and the most helpful resources in coping with the changes of 2011. These have been mentioned above for Brenda, John and Kevin.

Adam believed that the gradual change in assessment was mainly attributable to discussions between the school’s Physics teachers. Significant events included addressing staff meetings about the changes and discussions with other staff members who were not Physics teachers.

Alina had mostly worked alone, sharing occasional discussions with a teacher friend and obtaining tasks from a Physics teacher in another school.

Colin, at the start of 2011, was of the opinion that “the system’s not broken so why do we need to fix it”. A “significant milestone” had been the requirement to justify grades to students. He described the process of putting “unconscious thoughts on to a piece of paper” as challenging, but “a good process to go through”. Colin learned from the hub group and other teachers in the school. The clarifying forum was one of the “biggest points where I did change”, when he learned that his marking, based on SACE Board marking experience, was too harsh and, on returning to school, re-marked some previous tasks.

David had developed faith in the use of social moderation for the investigation tasks but not for the skills and applications tasks (SATs). He described changes he had made as “gradual”, with the discussions with other supervisors important because he could “make sure that what I was thinking matched up with what people I trust were thinking”. Moderation was “significant” because, although comfortable with his

effort, he was not comfortable with the reliability of the process.

Elena was satisfied her intra-school collaboration had been effective and made no acknowledgment of outside assistance.

For Frances, the most noteworthy event had been “lack of sleep”, a consequence of the additional time required for the use of performance standards. Although not alone in her school, Frances had effectively worked alone, with the others prepared to use what she produced.

Greg’s greatest resource was the hub group. It appeared, however, that Greg’s most significant “learning” had been at moderation where he became aware of the benefit of being a moderator and “in the know”.

Harry acknowledged that his beliefs had changed and described the change as “painful but useful”. Social moderation required him to design more rigorous investigation tasks, which subsequently influenced the tasks set for younger students. His major resource had been previous tasks that he “reworked”, with his IB experience being beneficial. The clarifying forum had stood out, where he discovered that his revised practicals were inappropriate, causing him to return to his earlier tasks, with minor adjustments.

Peter worked alone and remained convinced that external assessment was preferable to SBA and that SBA should be statistically moderated. He had, however, become receptive to using performance standards in designing tasks and in marking investigations. He also gave students more feedback than previously and began teaching students manipulative and report-writing skills.

Russell reflected how he, at the beginning of the year, “didn’t have a clue how I was going to use the performance standards”. He had used specific features in setting tasks and used a spreadsheet to collate data, a laborious task. Over the year he had gradually become more comfortable with the performance standards when he found they gave similar results to marking. He was more comfortable with marks and subsequently intended to “pay homage” to performance standards, using them in task design but not in grading. For Russell the most important learning event was after moderation when all moderators in the school reviewed their experiences and

discussed school practices for the following year.

Scrutiny of transcripts identified 74 change sequences involving external stimuli, as shown in Table 40.

**Table 40: Frequency of external stimuli in change sequences**

<b>External stimulus</b>	<b>Frequency</b>
Syllabus	21
Performance standards	18
Other teachers	11
Quality management processes	11
Exemplars	5
SACE procedures	5
IB processes	2
SACE Board officers	1

Table 40 shows that participants made most use (21 instances) of the assessment specifications of the new syllabus (SACE Board of South Australia, 2011d) and the performance standards (eighteen instances). Between them, they account for over half of the change sequences. This means that participants were active learners, using SACE Board documents to revise or generate their own assessment tasks and develop marking procedures. Other teachers were important stimuli (eleven instances), for example, the sharing of spreadsheets for managing students' grades. The quality management processes (eleven instances), most frequently the clarifying forums, initiated participant actions, such as re-marking previous work or revisions of task design. Thus, the quality management processes provided the external stimulus for 15% of the recorded change sequences.

Discussion with other Physics teachers was the most acknowledged resource for participants in this period of change. Where another teacher in the same school was present, collaboration was more valued, presumably because of greater discussion frequency. The importance of collaboration reflects literature reports of its importance in teacher learning (Fullan, 2008; OECD, 2009; Rosenholtz, 1989; Timperley et al., 2007; Wei et al., 2009).

Six participants cited SACE Board quality management processes as important resources: one mentioned the planning forum, three the clarifying forum and four

mentioned moderation. The references to these processes, however, were not always favourable, with the flawed tasks of the clarifying forum and awareness of “secret moderator business” viewed negatively.

It appeared that, although the quality management processes had stimulated some participant reflection on assessment practices, both the mandated syllabus changes and working with other teachers had been more important. For all participants, learning was strongly connected with reflection on their own practice, thus demonstrating experiential learning (Kolb, 1984) and reflective practice (Schon, 1987). Because participants’ reflections on their implementations of the changes were critical in their learning, the ongoing learning was in the hands of the teachers. Given that other factors—such as isolation, lack of moderation experience and individual motivation—influence teacher learning, it is questionable whether leaving teacher learning in the individuals’ domains is appropriate in a high-stakes environment. With most teachers working alone, common understandings are unlikely to arise. It is argued (see 8.4) that, with modifications, the quality management cycle should be used to enhance the learning opportunities for all teachers.

### **7.3.5 Overview of teacher learning**

Because the participants had developed new procedures for designing tasks and grading student work, it might be reasoned that learning had occurred. The participants described cycles of experimentation and reflection: experiential learning (Kolb, 1984). Each had, however, developed a unique set of procedures. There was less variation with tests than with investigations, possibly because the SACE Board had stated that the examination was unchanged so that participants felt less need for change. There were various ways in which participants had responded to the new assessment system. Differences in experience, knowledge and skills appeared to have generated diverse responses to the changes. The quality management processes contributed resources that participants used. The diverse responses suggest, however, that the SACE Board quality management processes were not enough to generate consistency among participants.

Common among all participants were changes in behaviour that were not acknowledged as improvements on the previous system, even when participants used



performance standards to provide better feedback (John) or as a faster means of marking (Kevin). No participant considered performance standards a “better way” (Evans, 2014) than the previous system. The argument that behavioural change, without recognition of a “better way”, does not constitute teacher learning (Evans, 2014), is discussed later (see 8.5.5).

Although participants acknowledged the quality management processes for elements of their learning, the processes should be modified to make them more effective contributors to teacher learning, as discussed later (see 8.4).

## **7.4 Teacher learning and teacher characteristics**

There was evidence of change in many aspects of the participants’ assessment practices. All participants adopted the specific features in the design of assessment tasks, making redundant any analysis of differences in task design practices in terms of teacher characteristics.

The eight participants who graded investigations using the wording of the performance standards included a mixture of teachers from different regions, different sectors and included both men and women. They included one person teaching SACE Physics for the first time and others with long experience. They seemed to represent a cross-section of the participants. Consequently, it was concluded that there was no evidence to suggest association between the adoption of change and characteristics of the participants, be they professional, school or personal. It is possible that engagement with the change process might be linked to personal characteristics that were not investigated in the current research, such as “perceptions of their own self-worth in the form of doubts, resistance, disillusionment, and guilt” (van den Berg, 2002, p. 599), as suggested later (see 9.4).

## **7.5 A summary of the evidence for teacher learning**

Overall, it appeared that the quality management processes contributed to teacher learning about the new assessment system. The most common stimuli for change were found in SACE Board documents. The exemplars, although criticised, were beneficial, and commonly used in writing tasks that made effective use of performance standards. The failure of the SACE Board to identify how specific

features might be assigned to tasks was criticised by all participants as unhelpful in their learning how to design appropriate assessment tasks using the performance standards. Opportunities for ongoing discussion, identified as important (e.g., Guskey, 2002; Saxe, Gearhart, & Nasir, 2001; Starkey et al., 2009; Timperley et al., 2007; Wei et al., 2010), were limited in the quality management processes. Feedback was reported to be problematic and unhelpful, both by those who produced it and the recipients.

These quality management processes appeared, however, to equip teachers to begin a learning journey—through awareness-raising and provision of resources—rather than sustain them during the journey. It is argued that if quality management processes are to nurture teacher learning, they need to be designed with teacher learning as an important goal and be more closely integrated into teachers' practices in their schools. In particular, it is desirable that teachers be informed of the rationale for change and be given opportunities to resolve any conflict with their existing beliefs. The next chapter discusses the findings and recommends changes to implementation programs and the quality management cycle that might facilitate teacher learning.

## CHAPTER 8: DISCUSSION

### 8.1 Purpose of research

The current research explored the potential for teacher learning in quality management processes associated with school-based assessment (SBA) in a high-stakes environment and examined the SACE Board processes in the light of the literature on effective learning for teachers and adults. Four aspects of the processes were investigated: teacher involvement, their form and content, and sought evidence of change in teachers' beliefs and practices in the area of assessment. The research contributes to the literature on the educative value of quality management processes, on teacher learning, on the use of social moderation in a high-stakes environment and on implementation of change in educational systems.

The research was conducted with Physics teachers during the implementation of changes in SACE SBA. Because the data emerging from the research consistently connected with the new expectations of teachers, further reading was undertaken on change in education systems. The process of ongoing engagement with the literature reflects grounded theory principles whereby the researcher is “open to all possible theoretical directions” (Charmaz, 2006, p. 46) and seeks to generate theory through an iterative process of engaging with the data and other relevant research. This additional engagement with the literature, near the end of data collection, provided reference points for data analysis and provided “an abstract explanation for the process being studied” (Creswell, 2008, p. 437), which is reported in this chapter.

Study of the quality management processes was compromised by conducting the research during the implementation of the assessment changes. Rather than separate examination of each research question, the discussion addresses threads that cut across the research questions and which extend beyond the initial purpose of the research. The discussion focusses on social moderation because that is the quality management process commonly addressed in the literature; other quality management processes are discussed where relevant.

It was found that research questions overlapped. For example, the number of teachers involved in moderation was related to the duration of that involvement,

linking the research questions on teacher involvement and the form of the quality management processes. The threads addressed include the number of teachers participating in social moderation, the duration of teacher involvement in social moderation, the potential for teacher learning in a quality management cycle and the potential for teacher learning in an induction program. The chapter concludes with a summary of recommendations arising from the research.

## **8.2 Number of teachers involved in social moderation**

The SACE Review Panel (Crafter et al., 2006) recommended greater involvement of teachers in moderation. This might imply participation of more teachers, an increase of time commitment, or both. Because the Review contrasted its recommended assessment system with the then-current system involving “a small group of designated moderators” (p. 130), it was concluded that the Panel envisaged an increase in the number of teachers involved.

The research demonstrated substantial involvement of teachers in the SACE Board activities associated with the 2011 assessment cycle that served two distinct purposes: quality management and induction. Quality management comprised “a four-phase cycle ... [of] planning, clarifying, confirming, and improving” (SACE Board of South Australia, 2010f), but which involved few teachers. In the new system three teachers serve as supervisors in each subject (SACE Board of South Australia, 2014j) and, in Physics, approximately 25 teachers serve as moderators. There was no substantial increase in the number of teachers involved in the quality management processes, as had been envisaged by the SACE Review.

If moderation is to be “an iterative process where shared understandings grow over time” (Hipkins & Robertson, 2011, p. 19), it is desirable that most teachers experience the discussion and negotiation (Adie et al., 2012; Klenowski & Wyatt-Smith, 2010b; QSA, 2009). Experts and novices should participate in moderation as both can make valued contributions and each can learn from the other (Hipkins & Robertson, 2011). Possible reasons why the new SACE failed to enact the vision of the Review Panel (Crafter et al., 2006) were examined.

### 8.2.1 Comparing the SACE and Queensland systems

The SACE Review Panel (Crafter et al., 2006) had been impressed by the district-based moderation of Queensland. The new SACE system, however, differs from Queensland's in ways that affect teacher participation. These differences include the apparently more collaborative arrangement between the authority and schools in Queensland than in South Australia, illustrated in a number of ways.

First, Queensland schools provide teachers for moderation without Queensland Studies Authority (QSA) funding (QSA, 2010a); only travel funding is provided.<sup>34</sup> In some South Australian schools there appears to be an expectation that the SACE Board fund teacher involvement. The current research revealed instances where schools had denied teachers approval for SACE panel involvement, suggesting that some school administrators do not perceive teacher learning as an outcome or do not value the learning, despite its being valued nationally (Australian Institute for Teaching and School Leadership, 2011a, 2011b).

Second, in Queensland there is no opportunity for separate moderation<sup>35</sup> of classes in a school: internal moderation is an expectation and a school responsibility (QSA, 2010e). In Queensland, schools are responsible for selecting the moderation sample as stipulated by QSA (QSA, 2010e), whereas, for the SACE, the Board stipulates the sample (SACE Board of South Australia, 2011f), which includes students from all classes. Furthermore, SACE moderators are directed to moderate work from all classes (SACE Board of South Australia, 2013d), so “checking” any intra-school moderation. Thus, even though intra-school moderation is recognised as an excellent vehicle for teacher learning (Hipkins & Robertson, 2011; Smaill, 2013), SACE schools may avoid it and are not trusted to undertake it.

A third difference lies in moderation outcomes. Queensland moderation panels make recommendations to the schools (QSA, 2010e), whereas SACE moderation panels

---

<sup>34</sup> Except where acknowledged, information relating to Queensland practices is based on information provided in personal communication by Natalie Carrigan, Quality Assurance Unit, Queensland Studies Authority, 13 July 2012.

<sup>35</sup> The QSA equivalent of social moderation is *verification*. For ease of comparison the term *moderation* is used in this document.

determine student grades (SACE Board of South Australia, 2014g) for SBA. A Queensland school may accept recommendations or negotiate with the district panel chair, which provides an opportunity for learning; negotiation is based on the match of evidence with the syllabus standards, not a compromise between the two parties. While South Australian schools may appeal after moderation (SACE Board of South Australia, 2014g), appeals relate to the process and not to the “results themselves” (SACE Board of South Australia, 2012d); no discussion occurs which might contribute to teacher learning. In Queensland, statistical adjustments are made to individual schools’ moderation outcomes for tertiary entrance purposes, with “each scaled against group results in the Queensland Core Skills Test” (QSA, 2010b, p. 15). In the SACE, no inter-school adjustments are made after moderation. The existence of such a post-moderation adjustment might help to address the concern expressed by some participants that teachers who are harsh markers are unlikely to have grades adjusted at moderation.<sup>36</sup>

Administrative differences also exist. Queensland’s moderation is district-based (QSA, 2010a), with thirteen districts (QSA, 2007b), whereas all SACE social moderation occurs in Adelaide (SACE Board of South Australia, 2011f). Queensland panellists are appointed for a six-year term (QSA, 2013) whereas SACE moderators are appointed annually. The longer QSA term may increase the likelihood of an “iterative process where shared understandings grow over time” (Hipkins & Robertson, 2011, p. 19).

It is not clear what factors contributed to the differences. In 2011, 358 Queensland schools submitted student work for Physics moderation by 208 panellists. In South Australia 25 panellists moderated 147 schools (SACE Board of South Australia, 2011a). In Queensland, 58% of schools provided Physics moderators, whereas in South Australia the proportion was 17%. It appeared that the changes to SACE assessment had not generated the greater involvement envisaged by the SACE Review Panel (Crafter et al., 2006).

---

<sup>36</sup> The statistical flags mentioned by SACE Board officers (see 8.5.3) might address this issue, but were not mentioned by any participants.

### 8.2.2 Why not the Queensland system?

While examination of Queensland's system was not a goal of this thesis, some comment is warranted on the failure to adopt Queensland's system. Possibly the South Australian political, educational and demographic environments were sufficiently different from those of Queensland that district-based panels were deemed inappropriate. From 1949 to 1988, South Australia experienced five inquiries into senior secondary assessment (Mercurio, 2003) but did not abolish external examinations as occurred in Queensland (Maxwell & Cumming, 2011). Possibly public dissatisfaction with examinations, as occurred in Queensland (E. Clarke, 1987; QSA, 2010c), was not experienced in South Australia. Unsuitability of district moderation for South Australia may arise from the greater proportion of its population in the capital; Queensland has a number of regional centres exceeding 50 000 (e.g., Cairns, Gladstone, Mackay, Rockhampton, Townsville) with South Australia having none. Other reasons, not in the public domain, may have arisen in the new SACE implementation. Regardless of the reasons, the SACE has fewer teachers involved in social moderation, with access to its claimed educative value.

The SACE Review Panel (Crafter et al., 2006) believed that the “overriding focus of assessment should be to help learners to become better learners” (p. 126) and reported “a gradual move away” (p. 126) from teacher involvement in moderation to statistical processes “that adjust teachers’ marks at a distance” (p. 126). The Panel argued that any cost of greater teacher involvement would “be returned many times over” (p. 130) through improvements in teaching and learning.

A consultation paper (*future* SACE Office, 2008) proposed two moderation models for SBA in the new SACE. The first model included intra-school and regional inter-school moderation with the acknowledgement that it offered “opportunities for more teachers ... to improve their professional capacity” (p. 21). The second model did not include intra- and inter-school moderation; thus the Panel’s vision (Crafter et al., 2006) had been downgraded to one of two options, rather than an integral part of the new SACE. For whatever reason, the second proposal was implemented. With the demise of the first proposal went opportunities for teacher learning, learning that is expected of Australian teachers (Australian Institute for Teaching and School Leadership, 2011b) and which principals are expected to support (Australian Institute

for Teaching and School Leadership, 2011a). An assessment system is less likely to deliver effective teacher learning if that learning is not valued and not considered to be a priority.

### **8.2.3 Moderation in a high-stakes environment**

Much research on the educative value of social moderation has been in middle and primary schools (Adie et al., 2012; Connolly et al., 2012; Klenowski & Wyatt-Smith, 2010a; Limbrick & Knight, 2005; L. Reid, 2007; Wyatt-Smith et al., 2010) and these are the levels where it is more commonly used. Its use in high-stakes senior secondary assessment appears to be limited to Queensland (QSA, 2009), ACT (ACT Board of Senior Secondary Studies, 2012), South Australia (SACE Board of South Australia, 2014i) and, previously, New Zealand (Hipkins, 2010a). Historically, group moderation<sup>37</sup> was used in some subjects in Victoria (Ingvarson, 1990) and South Australia (B. Atherton, 2009; Gipps, 1994a).

Because of perceptions that SBA is less reliable than examinations (Black, 1998; Crisp, 2010a; Harlen, 2005; Willingham et al., 2002; Wyatt-Smith, 1999; Wyatt-Smith et al., 2010), SBA is subjected to intense public scrutiny, particularly when used in high-stakes environments (Crooks et al., 1996; Maxwell, 2006). The number of teachers engaged bring any perceived inconsistencies of moderation (Hipkins, 2010a) “into the daylight” (Hipkins, 2010b, p. 18), whereas, in traditional examinations, inconsistencies are “hidden from ... all except the markers” (Hipkins, 2010a, p. 13; 2010b, p. 18).

To assure the public of the quality of their assessments (Mercurio, 2008), different systems respond to public scrutiny in any of five ways: using statistical moderation, providing professional development for moderators, providing post-moderation feedback, using specialist moderators and using controlled assessments.

#### **8.2.3.1 Statistical moderation**

Statistical moderation adjusts schools’ results statistically, usually against examinations (Matters, 2006), with no sighting of student work. It is used in most

---

<sup>37</sup> Group moderation in this thesis refers to confirmation or adjustment of student grades by consensus, with teachers involved in review of their own students’ work.



Australian states (Matters, 2006) and was used in the SACE before 2011 (SACE Board of South Australia, 2010j). The SACE Review (Crafter et al., 2006) recommended that no form of assessment be used to “statistically moderate another” (p. 133), arguing that “assessment modes should stand in their own right ... [and] that no assessment mode should dominate or be privileged” (p. 132). Questions have, however, been asked about “the legitimacy of the assumption that tests are inherently more reliable than teachers’ internal assessments” (S. Johnson, 2011, p. 42).

Participants in the current research spoke of greater reliability of examination marking than using performance standards, but also questioned the validity of examinations being used to moderate dissimilar tasks, such as practical work. Participants also mentioned that, without the requirement for student work to be seen, some Physics teachers did not undertake practical work. Colin contrasted the benefit of moderator discussion with its absence in statistical moderation. Statistical moderation is unable to provide teacher-learning opportunities and “can whittle away the professional skills of the teacher to design the assessment and make appropriate judgments” (Mercurio, 2008, p. 9).

Moderation that incorporates inspection of student work ensures that valid SBA has been undertaken, which is not guaranteed with statistical moderation, and can thus serve an educative purpose for teachers. It is recommended, therefore, that social moderation be continued for the quality control of SBA in the SACE.

### **8.2.3.2 Professional development of moderators**

Moderation is a “learning activity” (Hipkins, 2010b, p. 19) because it involves teacher conversations, which “inform assessment practice” (Hipkins, 2010a, p. 32). Consistency of teacher judgements in the New Zealand National Certificate of Educational Achievement (NCEA) was reported as taking several years (Hipkins & Robertson, 2011), involving “revised interpretations” (Hipkins, 2010b, p. 19) as understanding evolved. Hipkins (2010a) also wrote of “trends to tighten up moderation” (p. 39) to address perceived inconsistencies, presumably through better professional development of moderators. Crisp (2013) argued that markers’ (or moderators’) judgements are relative (citing Laming, 2004) and vary between people because of different experiences. If moderation, by broadening teachers’ experiences beyond their own students, is likely to improve the reliability of their judgements, it

would be beneficial that teacher participation be maximised.

In the current research most moderators believed that consistency had been developed between moderators, but that this might be improved with more discussion and more samples of student work. It is recommended that more teachers be employed as moderators and that professional development of moderators be enhanced to improve consistency in their judgements.

### **8.2.3.3 Post-moderation feedback**

“Feedback is one of the most powerful influences on learning” (Hattie & Timperley, 2007, p. 1), although it can vary in effectiveness. It is most useful when addressing faulty interpretations and providing advice on improving effectiveness (Hattie & Timperley, 2007), both of which might be expected in moderation feedback. In New Zealand, when post-moderation feedback was criticised for its variability, steps were taken to improve it (Hipkins, 2010a). Although the steps taken were not specified, later feedback was described as “more helpful” (Hipkins, 2010a, p. 4).

Post-moderation feedback is provided to Queensland schools “in supportive terms in the spirit of cooperation and professional respect” (QSA, 2010e, p. 30). The advice offered to moderators makes it apparent that freeform textual responses are included in the feedback.

In the current research, moderators and teachers were disappointed at the limited feedback provided after moderation. SACE Board officers described the feedback as “sanitised”. It appeared that the SACE Board preferred to limit feedback, being described as “hesitant to write anything down for fear it gets used against them”, rather than developing moderators’ capacities to provide more relevant and valued feedback. If teacher learning is a priority of social moderation, it is essential that informative moderation feedback be provided to teachers.

### **8.2.3.4 Specialist moderators**

Moderation involving professional judgement by one person is called *judgemental moderation* (the term is borrowed from Keeves, 1994) in this thesis. Johnson (2011) used the term *moderation by inspection* and described it as a “preferred strategy ... because the procedures involved are transparent and can be easily understood” (p.

42). The specialist moderators may be practising teachers (S. Johnson, 2011; SACE Board of South Australia, 2010j) or full-time moderators (Hipkins, 2010a).

The SACE Review (Crafter et al., 2006) decried moderation processes that placed “authority and responsibility in the hands of a small group of designated moderators” (p. 130) and recommended involving more teachers because of the “benefit to the teachers involved, and their students” (p. 130). Previous research has reported the educative value of moderation experience (B. Atherton, 2009; Barrett, 2008; Davidson, 2003; Ingvarson, 1990; Limbrick & Knight, 2005; L. Reid, 2007). Although the new SACE employs a small group of moderators, moderation entails discussion. In the current research, moderator-participants reported grading consistency and improved personal assessment knowledge and confidence. Consequently, it is recommended that more teachers be given this opportunity and that moderators be recruited solely from practising teachers, unlike the current system in which some retired teachers also moderate.

#### **8.2.3.5 Controlled assessment**

Because of concerns about the validity and reliability of SBA, it was replaced in the UK General Certificates of Secondary Education by *controlled assessments* (Crisp, 2013) that are set, conducted and marked under the awarding body’s conditions (QCA, 2008). Research was undertaken on teachers’ responses to their introduction (Crisp & Green, 2013). For assessing important skills, 43% of teachers believed there was no difference between SBA and controlled assessment, while 34% preferred SBA and 24% controlled assessment (Crisp & Green, 2013). Similar proportions were reported for their encouraging learning of important skills. Two thirds of teachers, however, reported that controlled assessment had negated the purpose of SBA, reducing its learning benefits, with tasks becoming “formulaic with too much structure provided by teachers” (Crisp & Green, 2013, p. 693).

Crisp and Green (2013) also reported that the potential for plagiarism had been a key factor in introducing controlled assessments and they stated that 57% of teachers believed it advantageous over SBA in combatting plagiarism. The inclination for students to cheat, however, depends on many factors (Kohn, 2007), including the student-teacher relationship, the nature of the task and the academic environment. It is important that, if SBA is to be valued, plagiarism needs to be addressed. There are

also other criticisms of SBA that need to be addressed, including potential teacher bias and poor reliability (Black, 1998; Harlen, 2005; Willingham et al., 2002; Wyatt-Smith, 1999; Wyatt-Smith et al., 2010). It is recommended, therefore, that ongoing quality management processes incorporate professional development for teachers to assure the quality of SBA, obviating the need for alternatives, such as controlled assessment with its negative effects on student learning.

### **8.2.3.6 Preferred moderation practice**

The selection of a preferred form of moderation appears to be based on its purpose and public perception. Examinations are held in esteem although this may be due to ignorance of examination practices (Hipkins, 2010a, 2010b) or for social reasons (Black, 1998). Criticism of external examinations (e.g., Baird, Greatorex, & Bell, 2004; Broadfoot, 2005; Gipps, 1995; Suto & Nádas, 2009) is not widely disseminated and knowledge of examinations has been described as a “dark alley” (Broadfoot, 2005, p. 129). It is recognised that most examinations are “amenable to coaching” (Gipps, 1995, p. 273) and that pupils with different school experiences are not equally prepared to compete. Discussion among markers is educative (Baird et al., 2004), which was mentioned by participants in the current research. Adam, for example, described the “vigorous” debate at markers’ meetings as “excellent” professional development. Teacher discussion has been recognised as an excellent form of teacher learning (Hipkins & Robertson, 2011; Loucks-Horsley et al., 2003; Wei et al., 2009; Zemke & Zemke, 1988) and involving more teachers in social moderation would disseminate knowledge more broadly than with a small panel of moderators or markers.

The QSA, in itemising the purposes of assessment, begins with “to help schools achieve quality learning outcomes for their students” (QSA, 2009) and declares that its assessment “above all ... promotes authentic pedagogy and confirms the role of teachers as professionals” (QSA, 2010c, p. 4). *The New Zealand Curriculum* (Ministry of Education, 2007) declares: “The primary purpose of assessment is to improve students’ learning and teachers’ teaching” (p. 39). Conversely, the Act defining the SACE Board’s role (*SACE Board of South Australia Act 1983*, 2010) focusses on assessment; the only mention of improving teaching and learning is that the Board must “take steps to assist schools ... to provide information and professional

development opportunities to teachers” (Section 15 (3)). This may explain why the Board did not see supporting teaching and learning as a priority. When approached at the beginning of the current research, the SACE Board Chief Executive<sup>38</sup> indicated that the primary purpose of moderation was to assure the quality of SBA, but acknowledged that involvement might also provide teacher learning opportunities. In a focus group, an officer explained that moderation was “not actually PD [i.e., professional development] because ... that’s not the intention of moderation, it is just a by-product.” In an online forum notice (SACE Board of South Australia, 2014h), however, teachers were advised that “One way teachers can further develop their understanding of performance standards and assessment is by nominating to be involved in marking and moderation panels.” It appears that the SACE Board uses the educative nature of moderation to attract panellists but does not prioritise strategies that might maximise its educative value. Failure to design quality management processes that do facilitate teacher learning is likely to compromise their potential for teacher learning. Teacher participation in social moderation might, indeed, be stimulated if it is promoted as one way of satisfying national expectations of professional learning for teachers (Australian Institute for Teaching and School Leadership, 2011a, 2011b).

The current research showed that the most common reason for participant involvement in the quality management processes was, in fact, the expectation of learning. Adult learning principles recognise relevance to the learners’ needs, congruence with prior experience and the use of appropriate activities and materials as central (Díaz-Maggioli, 2004; Knowles et al., 2005; Loucks-Horsley et al., 2003; Zemke & Zemke, 1988; Zepeda et al., 2013). Participation in social moderation satisfies these requirements for effective teacher learning. When improved teaching and learning are seen as important outcomes, authorities are more likely to implement processes that provide teachers with such opportunities.

It is recommended that more practising teachers participate in moderation and that well-designed professional development be provided to enhance the reliability of moderation decisions. Because standards emerge gradually (Hipkins & Robertson,

---

<sup>38</sup> Private email correspondence.

2011), it would appear that, over a longer time, greater consistency would result and would be facilitated with continuity of moderators. It is, therefore, recommended that moderators be appointed for a period greater than one year. Since it is unlikely that all subject teachers can serve as moderators, it is also recommended that moderators be afforded opportunities to develop their skills in writing feedback informing teachers of how their SBA may be improved.

### **8.3 Duration of teacher involvement in social moderation**

The current research found that approximately 15% of SACE Physics teachers experienced the necessary duration for effective professional development (Desimone et al., 2002; Garet et al., 1999; Garet et al., 2001; Porter et al., 2000). There is, however, a conflict between the number of teachers involved and the duration of the involvement: employing more moderators requires less time to complete moderation. Queensland engages more teachers for a shorter time; South Australia engages fewer teachers for a greater time.

It could be reasoned that the smaller number of SACE moderators derive a greater benefit from their experience than those in Queensland since they undertake approximately twice as many hours each year. The price paid is that fewer teachers have the opportunity of that experience. It appears that participation in social moderation alone is unable to satisfy the need for sufficient duration of engagement for enough teachers. The next section discusses possibilities of other quality management processes addressing this problem.

### **8.4 Teacher learning from the quality management cycle**

Although social moderation has been advocated for its educative benefit to moderators (Harlen, 2004a; Hipkins & Robertson, 2011; Klenowski & Wyatt-Smith, 2010b; Maxwell, 2006), the current research differed from previous research into social moderation (e.g., Adie et al., 2012; Connolly et al., 2012; Davison, 2004; Hipkins & Robertson, 2011; Klenowski & Wyatt-Smith, 2010a) by also investigating other quality management processes that engage teachers annually.

Performance standards are open to interpretation (Crafter et al., 2006; Klenowski & Wyatt-Smith, 2010b) and teachers need discussion opportunities to develop common

understanding (Adie, 2008; Crafter et al., 2006), which requires time (Davison, 2004; Fastier, 2007; Hipkins & Robertson, 2011). The participants in the current research were not representative of SACE Physics teachers (see 4.1.1.4 and 4.2.5), with the high proportion with SACE panel involvement being noticeable. Even so, five participants appeared unlikely to experience moderation in the next few years after the research concluded, with inexperience and school commitments the major reasons. Among SACE Physics teachers as a whole, it is possible that lack of interest might contribute to a higher proportion without moderation experience.

In the current research, teachers believed they had been provided with insufficient advice on how performance standards could be used in task design and in grading of student work. The current research provided evidence that the other quality management processes might also contribute to teacher learning, that the annual quality management cycle could be used to improve teachers' task design and grading skills. This section argues that the ongoing practice in these processes would enable greater teacher involvement, thus contributing to teacher professional learning, and examines the potential of intra- and inter-school collaboration, the use of feedback and exemplar tasks.

#### **8.4.1 Intra-school collaboration**

This research found that the most effective communities of practice were the collaborations with other teachers in the same school, supporting literature reports of its importance in teacher learning (Hawley & Valli, 1999; Timperley et al., 2007; Wei et al., 2010). Inter-school collaborations were effective but of limited duration, falling into disuse after the first year; the complexities of meeting with teachers in other schools was mentioned by participants. It appeared the inter-school collaborations had been established for teachers to “compare notes” and share resources but that, after the first year of the new system, the need had passed. Suggestions for facilitating collaboration outside a teacher's school are discussed in 8.4.2.

Because conversations about the application of standards to student work can “inform assessment practice” (Hipkins, 2010a, p. 32), intra-school moderation would be beneficial (Adie, 2008; Connolly et al., 2012; Daugherty, 1997; Wyatt-Smith et al., 2010) and should be encouraged. Six participants described discussing,

negotiating and sharing assessment tasks with colleagues in “a collective process of negotiation that reflects ... mutual accountability” (Wenger, 1999, pp. 77, 78).

Except for one participant, who was new to the school, the collaborations continued practices that existed before 2011. Intra-school collaboration appeared more effective than inter-school ones, with implications for teachers working in isolation, addressed below.

Hipkins and Robertson (2011) referred to moderation as “learning together” (p. 17) and Smaill (2013) argued the need for “regular opportunities” (p. 259) for teachers to participate in social moderation. It would appear that the simplest means for generating such opportunities would be in schools with more than one teacher of the same subject. The revised SACE assessment system, however, rather than encouraging such practices, allowed schools to avoid them. An assessment system with teacher learning as a priority would require schools with multiple classes in a subject to submit as one group; schools, rather than the SACE Board, would have the responsibility for ensuring comparability of results across classes. It is recommended that, regardless of the number of classes, all students in a school comprise one subject assessment group and that the school has responsibility for intra-school moderation. This is more likely if teacher learning is seen as a priority of the assessment system rather than an incidental outcome.

#### **8.4.2 Extra-school collaboration**

While participants reported having other Physics teachers in the school to be beneficial for teacher learning, five participants were the only Physics teachers in their schools. This is not uncommon in South Australia where two-thirds of the Physics teachers are the sole teacher of the subject in their school (SACE Board of South Australia, 2011a). To provide ongoing discussion opportunities (Daugherty, 1997; Guskey, 2002; Saxe et al., 2001; Timperley et al., 2007; Wei et al., 2010), inter-school collaboration is necessary if all teachers are to have the opportunity of the benefit of professional discussion (Adie, 2008; Hipkins, 2010a; Hipkins & Robertson, 2011; Ingvarson, 1990; Limbrick & Knight, 2005; L. Reid, 2007; Timperley et al., 2007; Wei et al., 2010). Pre-moderation advice and online collaboration are examined as means of facilitating teacher access to professional discussions.



### **8.4.2.1 Pre-moderation advice**

In New Zealand, “pre-moderation of assessment tasks” was one means of assuring the reliability of NCEA assessment (Ministry of Education, 2011, p. 32) and “pre-assessment moderation” was reported to help teachers to clarify their NCEA assessment understandings (Hipkins & Robertson, 2011, p. 6). Participants in the current research, recognising the importance of task design, lamented the lack of opportunity for having tasks checked before their use with students. A few participants created opportunities for sharing their tasks with teachers in other schools. Although no participants submitted tasks to SACE Board officers, officers reported teachers approaching them for advice, but were unable to satisfy all requests, indicating the need for a more effective supportive mechanism, particularly early in the innovation. In the SACE, supervisors define interpretations of the performance standards and in the current research supervisor-participants reported how these understandings evolved. Performance standards provide a basis for teacher knowledge and understanding (Adie, 2008) and these will change with further experiences, because teachers’ interpretations are influenced by their backgrounds and experiences (Burt & Stapleton, 2010): understandings of performance standards are not fixed, but evolve. Consequently, supportive mechanisms should not be restricted to the induction phase of an assessment system but must be maintained further. They are part of the totality, not merely part of the induction.

Prior to the new SACE, some subjects had *support moderators* (SSABSA, 2006) providing advice to teachers in task design and grading of student work throughout the year. Their benefit was acknowledged previously (B. Atherton, 2009) and in the current research by participants and SACE Board officers, but they were not included in either of the models proposed for the new SACE (*future SACE Office*, 2008). Support moderators could advise to teachers on their task design and grading decisions during the year, so they may improve their practices prior to the end-of-year moderation. Provision of extensive pre-moderation advice by SACE Board officers appears to require more officers, with budgetary implications.

It is recommended that the SACE Board should cease payments to schools for releasing teachers for moderation and that the money be used to provide support

moderators. It is suggested that, where schools need to engage replacement teachers, funding be from schools' professional development budgets. Instead of moderation being seen as "SACE business", it needs to be seen as teacher learning, which is "school business" (Australian Institute for Teaching and School Leadership, 2011a, 2011b). This paradigm shift would be more likely to have been accepted as an integral part of the new SACE, since the SACE Review (Crafter et al., 2006) endorsed the educative value of moderation. Support moderators offer a proven means of facilitating teacher learning and their funding is more likely if the assessment system acknowledges teacher learning as an important outcome of SBA.

#### **8.4.2.2 On-line collaboration**

During 2011 the Physics online forum experienced increased activity. Although subject teacher associations provide online discussion groups, they are restricted to association members. In the current research, offers of alternative forums for sharing generated little response. The most appropriate forum for Physics teacher discussion appears to be the SACE Board online forum, since most Physics teachers are already members. Although it could be argued that teachers bear responsibility for stimulating online discussion, it was found that this was limited. If teacher learning is a priority of the assessment system, SACE Board stimulation of online discussion is one means of facilitating such learning.

One purpose of the online forum might be providing online access to student work. In the current research, such access was provided prior to the clarifying forums. During implementation of the new SACE, the possibility of online moderation was explored (SACE Principals Partnership Strategy Group, 2009), although an internet search failed to find any report on the matter. Sharing of teachers' opinions online might generate the discussion and negotiation necessary for the development of shared understandings (Adie et al., 2012; Hipkins & Robertson, 2011; Klenowski & Wyatt-Smith, 2010b; QSA, 2009; Wyatt-Smith et al., 2010). For many, the only teacher of the subject in a school, this might be the only opportunity for such discussion. A search of the Education Resources Information Centre (ERIC) database revealed limited literature on the possibility of online moderation, with one article (Adie, 2013) reporting that online discussion enhanced teacher understanding; teachers acknowledged new perspectives developed in discussions with "others

outside of their local context” (p. 105). Other papers dealing with online moderation were also identified (Adie, 2008, 2011). Participants reported the use of online methods in the International Baccalaureate’s (IB) professional development for moderators. If teacher learning is to be encouraged as a system priority, online sharing of student work is one means by which this might be achieved.

The role of Chief Supervisor—Materials Development might be used to stimulate online discussion of assessment tasks and student responses among teachers. There are currently eight expectations of this role (SACE Board of South Australia, 2014j); it is suggested that stimulation of online discussion might be added. This could be as little as providing one item each term, for example learning and assessment plans (LAPs), assessment tasks and student work. Initially, anonymous tasks could be as presented to stimulate discussion on task design and the interpretation of specific features. Later, completed tasks could stimulate discussion on interpretation of the performance standards. Protocols would need to be established and some teachers might prefer anonymous discussion contributions.

Teacher participation cannot be assumed. The time required would, however, be less than a day or half-day away from school and could be done at teacher convenience. It would provide sole subject teachers with an opportunity for professional dialogue not otherwise readily available. Teachers could be invited to offer tasks for discussion under condition of anonymity, if preferred. The Chief Supervisor—Materials Development could contact members of the SACE Board’s Curriculum Leaders Group<sup>39</sup> for the subject, requesting comments on posted tasks. After online discussion, the Chief Supervisor could summarise and offer a revised version to the author, or the revised task could be added to the website.

In this way, the system could offer teachers a growing body of assessment tasks, thus assisting in clarification of “the intent and expected achievement levels of tasks ... provided” (Hipkins & Robertson, 2011, p. 6). The availability of exemplars and “guidance ... in providing appropriate reference points” (Crisp, 2013, p. 130) would

---

<sup>39</sup> Curriculum Leaders Groups provide expert subject teacher advice to the SACE Board, including contributing to the professional learning of teachers and to the development of subject-specific support materials (SACE Board of South Australia, 2014c).

assist teachers in their assessment practices. If teacher learning is a system priority, numerous examples would assist the development of a common understanding among teachers, not possible when each teacher's experience is restricted.

### **8.4.3 Feedback**

Although the power of feedback in learning has been demonstrated (Hattie, 2009; Hattie & Timperley, 2007), its effectiveness varies (Bransford et al., 1999; Fullan, 2011; Gielen, Peeters, Dochy, Onghena, & Katrien Struyven, 2010; Hattie & Timperley, 2007; Hipkins, 2010a), being most effective when it occurs regularly and in time for implementation (Bransford et al., 1999; Gibbs & Simpson, 2004). It also needs to be embedded in a supportive learning culture (Fullan, 2011), which is expected to be developed by school principals (Australian Institute for Teaching and School Leadership, 2011a) and is desirable in SACE Board processes.

Participants in the current research were critical of moderation feedback, reflecting a number of these limitations. Feedback is received after students' results have been submitted and does not appear to occur in a learning-supportive climate, since, as SACE Board officers explained, professional development "is not the intention of moderation". The feedback should focus on understanding (Bransford et al., 1999) and needs to provide specific information upon which the learner can act (Black & Wiliam, 1998; Hattie & Timperley, 2007). Lack of specific information in feedback was severely criticised by participants.

If teacher learning is a system priority, the provision of informative school-specific feedback is essential. It is recommended that the nature of feedback provided in Queensland and New Zealand be explored and that the professional development used in those places be considered to equip moderators to provide more informative feedback to teachers of the SACE.

### **8.4.4 Exemplars**

Exemplars have been identified as important elements in teacher learning (Harlen, 1994; Maxwell, 2006; Sadler, 1987). Crisp (2010a) advocated their use in SBA to "provide assessors with reference points against which comparisons can be made and hence can help to fix the standard" (p. 53). Sadler argued that exemplars assist teachers to make explicit criteria and standards that exist "in an unarticulated form"

(1998, p. 81) and that many exemplars are needed when there are many criteria (1987). The need for multiple exemplars was endorsed by others (Crisp, 2013; Hipkins & Hodgen, 2011; Klenowski & Wyatt-Smith, 2010b; Stanley et al., 2009).

Participants in the current research used the SACE Board exemplars in different ways, mostly for assistance in assessment task design, and a majority were appreciative. There were, however, criticisms, with both the small number and quality of exemplars; most participants wanted more exemplars, with greater variety of task type and with specific features assigned. Greg believed that “the best way of learning about good task development is to just see lots of good tasks”.

Stanley et al. (2009) argued that exemplars need to be “electronically available and regularly augmented and refreshed” (p. 80). It was considered that, over the last four years, the exemplars have been supplemented and updated only to a minor extent. At the end of the research period, participants were still critical of insufficient exemplars and some had ceased referring to the website because no new ones were being provided. It is recommended that, with teacher learning as a priority, website exemplars be supplemented and that teachers be engaged in the development of such materials as described above.

#### **8.4.5 Ongoing teacher learning**

In the current research, it was found that correct use of the term *performance standard* had increased but there was little evidence that incorrect usage had declined. There was questionable evidence of improved consistency in applying performance standards in the grading of student work. Familiarity with terminology and shared understandings of performance standards arise with their use in a social context (Adie et al., 2012; Bransford, Brown, & Cocking, 2000; Eraut, 2007; Hipkins & Robertson, 2011; Klenowski et al., 2007; Lave & Wenger, 1991; Wyatt-Smith et al., 2010). Teachers working alone, the majority of Physics teachers, are denied the opportunity of increasing their understandings if they are not required to articulate those understandings. Even the small proportion of teachers who serve as moderators, approximately 15%, experience only a few days of the year in which articulation of their understandings is necessary. It is recommended that the ongoing quality management cycle be used to provide opportunities throughout the year for teachers to increase their understandings of assessment and the use of performance

standards in task design and in grading student work.

## **8.5 Teacher learning from induction programs**

In the current research, processes initially considered to be part of the ongoing SACE Board quality management cycle (SACE Board of South Australia, 2011e), were later determined merely to be part of the induction of teachers to the new SACE. This section concerns those processes that, being one-off for most teachers, were considered to be part of the induction program and discusses how teacher learning might have been facilitated.

There were many positive aspects of the SACE Board induction program, demonstrating many features recommended in the literature on high quality professional learning: active learning, coherence with previous and subsequent practice, a subject focus and collective participation by teachers of the same subject. The following sections examine aspects of effective teacher learning that were not so evident in the induction program—providing a rationale for change, challenging teachers’ preconceptions, teacher involvement in planning and the duration of the implementation—and examines the argument that teacher learning requires more than behavioural change.

### **8.5.1 Providing a rationale for change**

The evidence indicated that the quality management processes failed to inform teachers adequately of the reasons for the changes. Adults are more motivated to learn if convinced of the benefit they derive (Donavant, 2009; Fogarty & Pete, 2004; Knowles et al., 2005) they derive. Induction of teachers to changes should focus on broad issues such as the principles behind the changes, rather than mere operational matters (Starkey et al., 2009). Teachers provided with a coherent rationale for a change in their practices would be more willing to commit to the change, whereas without sufficient rationale, acceptance of the change is less likely and the changes might be met with resentment and a lack of motivation (Donavant, 2009).

Providing adequate reasons for imposed change, and discussion opportunities for teachers to internalise the rationale, should be an essential part of teacher induction programs.

### 8.5.2 Challenging teacher preconceptions

The induction program could be criticised for failing to challenge teachers' existing understandings where they conflicted with the changes being introduced (Loucks-Horsley et al., 2003; Thompson & Zeuli, 1999; Timperley et al., 2007). Participants recalled reassurances that nothing much had changed and SACE Board officers described the role of the planning forums as reassuring teachers that "it's not going to be a scary thing". Participants also reported that some forum discussions failed to resolve issues raised, although they acknowledged that this was possibly due to lack of time.

The effort people put into learning is influenced by learners seeing the usefulness of that learning (Bransford et al., 1999; Knowles et al., 2005) and most teachers participate in professional learning activities "to become better teachers" (Guskey, 1986, p. 6). Timperley et al. (2007) pointed out that teachers need powerful reasons "to engage with new information in sufficient depth to change their practice" (p. xxxviii). It was considered that failure to challenge teachers encouraged a surface level approach of compliance, rather than a deeper level of engagement. Hargreaves et al. (2001) contend that change is complex and that changes in classroom assessment represent "major paradigm shifts [requiring] ... rethinking what assessment and teaching are for" (p. 50). Successful implementation of change requires teachers "committed to the goals of the reform process" (Gitlin & Margonis, 1995, p. 378), which depends on their being convinced the change is "feasible and will benefit their students" (Hargreaves et al., 2001, p. 118). A deeper level approach, with greater teacher commitment to change, is more likely if professional development focusses on the principles underlying the changes rather than simply teaching the necessary operational skills (Starkey et al., 2009, citing Black). Effective implementation is more likely if teachers are persuaded to "think differently about their behavior, prompting them to raise questions about their existing behavior" (Spillane et al., 2002, p. 421). Teachers have "a self-reinforcing belief ... [that assists them] in maintaining, rather than changing, their existing practice" (Yerrick et al., 1997, p. 140); they are able to assimilate innovations without changing their fundamental beliefs. In the current research it was found that, at the end of the research period, after working with performance standards for twelve months, all participants remained unconvinced of their benefits. Because

acceptance of change is more likely if teachers are convinced of its benefit to students, it is recommended that induction programs be designed to encourage deeper understanding, with time for resolution of any dissonance.

### 8.5.3 Teacher participation in induction planning

Learner involvement in planning has been recommended for adult and teacher learning (Knowles et al., 2005; Loucks-Horsley et al., 2003). There is also recognition that, in planning teacher learning, attention be given to learners' expectations and prior experiences (Díaz-Maggioli, 2004; Donavant, 2009; Zemke & Zemke, 1988, 1995).

SACE Board officers reported that the adoption of a common format for all subjects reduced the effectiveness of some parts of the planning forums. While teachers of some subjects, such as Physics, were interested in the “statistical flags”<sup>40</sup>, others were not. Greater subject teacher involvement in planning of professional development activities is recommended, thereby informing facilitators of subject teacher prior knowledge and opinions on the changes being introduced, which would allow informed subject-specific modifications of the induction program.

### 8.5.4 Duration of an implementation program

The SACE Review Panel acknowledged that senior secondary students “need time to develop a range of increasingly sophisticated ways of thinking and working” (Crafter et al., 2006, p. 126) and teachers need time to implement change (Evans, 2014; Guskey, 2000; Hipkins & Robertson, 2011; Loucks-Horsley et al., 2003; Wylie & Hodgen, 2010). Hiebert (reported in Garet et al., 2001) said that the time needs to be measured in years and one to two years has been suggested (Timperley, 2008), although “several years” were required to develop judgement consistency in the New Zealand NCEA (Hipkins & Robertson, 2011). Three years was needed in UK experience (Daugherty, 1995) and three to five years has also been mentioned (Loucks-Horsley et al., 2003). The “progressive development of deeper knowledge

---

<sup>40</sup> A SACE Board panel uses statistical processes to compare the results of the external and SBA components for each school. Where moderation does not resolve discrepancies, a second moderation panel decides whether the SBA grades should stand, or whether they need to be adjusted (SACE Board of South Australia, 2010c).



and understanding” (Adie, 2008, p. 2) requires time for reflection (Bissaker, 2009; Bransford et al., 2000; Cross, 2007; Eraut, 2007; Hoekstra & Korthagen, 2011; Moon, 2000; Pollard, 2002; Schon, 1987; Shulman, 2002). If “positive change involves continuous learning” (Hargreaves et al., 2001, p. 118), implementation is not a short-term event. Thus it is not unexpected that one-off sessions—the old paradigm—are contrasted with the “new”, which recognises that “lasting change will be a slow process” (Hawley & Valli, 1999, p. 134). Teacher skill development requires “both time and effort” (Guskey, 2002, p. 386) and is facilitated by ongoing practice with follow-up and support rather than one-off events.

The current research supported the picture of progressive development, with all participants describing a “fairly gradual” process, with no “sudden changes”. Adam, for example, spoke of a “growing understanding”.

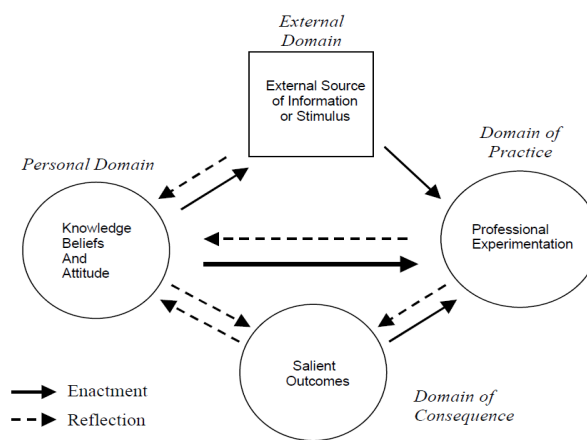
Although policymakers’ proposals help teachers to develop new understandings, implementation can also generate new understandings (Spillane et al., 2002). “Policy evolves as it is implemented” (Spillane et al., 2002, p. 419) and the gradual learning observed in the current research applied to the SACE Board as well as the teachers. One SACE Board officer described how, at the time of the planning forums, many details had not been decided and consequently some questions could not be answered. Supervisor-participants described evolutionary learning processes as “trying to stay a step or two ahead” and “making it up as we go”. The comment “It’s hard to ... pick a standard at the beginning of the year” illustrates the way in which standards evolve over time (Hipkins, 2010b; Hipkins & Robertson, 2011); the three supervisors, having agreed among themselves of the standard of a student task, later revised their opinions. If policies that teachers are required to implement mutate during implementation, teachers require familiarisation with changing expectations, necessitating ongoing induction.

Induction needs to be prolonged, allowing time for reflection (Cross, 2007; Eraut, 2007; Loucks-Horsley et al., 2003; Moon, 2000; Pedder et al., 2005; Schon, 1987) in a supportive environment. The Assessment Reform Group (ARG, 2006) identified both pre-service and in-service professional development as necessary for the implementation of SBA. In the current research, teacher learning could best be described as experiential, involving sequences of active experimentation followed by

reflection and modification. It is, therefore, recommended that induction programs be implemented over a prolonged period, not restricted to introducing change, and that programs be designed with the intention of facilitating and enhancing teachers' experiential learning.

### 8.5.5 Learning and mandated change

The current research has viewed teacher change through the interconnected model of teacher change (D. J. Clarke & Hollingsworth, 2002), shown in Figure 5, repeated below.



**Figure 19: Clarke and Hollingsworth's interconnected model of teacher learning**

The new syllabus and the quality management processes provided external stimuli for teacher change. Participants' descriptions of learning during the year satisfied the model's construct as occurring via reflections and enactments—experiential learning (Kolb, 1984). There were rare instances of participants valuing an imposed change, such as Adam's appreciation of the need for a LAP, supporting the view that changes in beliefs can follow changes in practice (Guskey, 1986). More commonly, reflection was in terms of the effectiveness of the implementation rather than any valuing of the change. Although, for example, John used performance standards in creating feedback for students, he did not value performance standards as such and, at the end of the research period, would prefer not to use them. Participants did not value the changes in assessment as a "better way" (Evans, 2014).

Evans (2014) argued that although changes in beliefs or attitudes are only adopted when viewed as a "better way", the same is not true of behavioural changes, which

may be imposed. She considered behavioural changes, without attitudinal change, as professional development rather than “professional learning, which requires intellectual or attitudinal development” (p. 192). Scrutiny of the changes described by participants, however, shows many instances of reflection on their professional experimentation that led to revision. It was considered that “intellectual development” was evident and that professional learning had occurred without acknowledging the change as a “better way”.

## **8.6 First Year Evaluation of the new SACE**

Before the final interviews in the current research, the SACE Board announced an evaluation to “identify and analyse the initial effects of the new arrangements for the SACE and to determine where it would be appropriate to implement improvements” (SACE Board of South Australia, 2012g, p. 2). The Report (SACE Board of South Australia, 2012b) was released during the research period but, to minimise influencing the current research analysis, reading was postponed until data analysis had been completed. Although the Evaluation made recommendations on many aspects of the new SACE, the analysis below is restricted to assessment and teacher learning, particularly performance standards, exemplars, assessment task design, moderation feedback and clarifying forums.

### **8.6.1 Performance standards**

The First Year Evaluation (SACE Board of South Australia, 2012b) noted that, if teacher judgement is to be valued, teacher understanding of the performance standards is essential. The Evaluation Panel reported receiving no criticism of the concept of performance standards, but many of teacher difficulty in their application, particularly in “the interpretation of adverbs and adjectives” (p. 79). This reflects participant comments that they were “very hard to get your head around” and were expressed in “pretty indigestible prose”. The Report called for greater clarity and proposed that clarifying forums be used to obtain feedback from teachers about how they might be made clearer. A difficulty with this suggestion lies in the minimal attendance of teachers at clarifying forums, compounded by attendance being facilitated for teachers experiencing less success with the assessment system; teachers with assessment success, potentially able to offer good suggestions, are less likely to attend.

### **8.6.2 Exemplars**

The Panel (SACE Board of South Australia, 2012b) was not surprised at the lack of sufficient and suitable exemplars in the first year of the new SACE and expressed concern at the effect on teacher “confidence in determining grade levels without adequate exemplars” (p. 27). In the current research, Adam lamented the difficulty of moderating grades for which no exemplars were available. The Panel recommended an increase in the number and range of annotated exemplars of student work for the various grades of SBA tasks. This supports comments made by research participants and reported in the literature (Crisp, 2013; Hipkins & Hodgen, 2011; Klenowski & Wyatt-Smith, 2010b; Stanley et al., 2009). The Panel reported that the process of reviewing and extending the exemplars had commenced. In the current research, inspection of the SACE website showed, however, that more work was needed on this aspect.

### **8.6.3 Assessment task design**

The First Year Evaluation (SACE Board of South Australia, 2012b) observed the importance of teacher expertise in “the design of assessment tasks which enable the students to demonstrate their learning consistent with the expectations of the performance standards” (p. 91) and reported the distress of some experienced teachers who had student work downgraded because of the task design. This fear of disadvantaging their students was expressed by many research participants. John, for example, worried “that I’m going to do my students a disservice”. At the back of Adam’s mind was the effect on “the kids if I bugger it up”.

The Panel, consequently, saw the importance of assisting teachers to improve the design of their assessment tasks, suggesting that this be done “in association with teacher associations and using the ... experience of practising teachers” (p. 13). Participants had commented that some subject teacher associations had generated helpful material but that the SA Science Teachers Association had not done so during the research period.

### **8.6.4 Moderation feedback**

The First Year Evaluation Report (SACE Board of South Australia, 2012b) noted, for an assessment system that values teacher judgement, the importance of continued

improvement in teacher understanding of the rationale for moderation changes. It reported the failure to provide feedback on task design and commented that this had “caused considerable anxiety for a number of teachers” (p. 80), particularly when task design flaws were identified as the reason for grade changes. The Evaluation Panel described teachers as “clamouring for more feedback.” (p. 86). This was also reported by SACE Board officers and participants in the current research, with feedback described as “disappointing” and “terrible” by teachers and moderators alike.

The Panel acknowledged the increased detail provided in Chief Examiners’ reports and supported their continuance and expansion. While the Evaluation report recommended “strengthening the feedback from moderation processes” (p. 6), it conceded the difficulty of “providing direct feedback to teachers about each of their assessment tasks as well as the reasons for the outcomes from the moderation process” (p. 91). The Panel recommended that the SACE Board authorise the provision of “expanded and direct (i.e., including face to face) feedback” (p. 14) to teachers with large moderation changes or where assessment task design had been the key determinant in the moderation outcome. It did not, however, suggest how to resolve “the difficulty, from a resource and logistics viewpoint” (p. 91) of providing such feedback.

### **8.6.5 Clarifying forums**

The Report (SACE Board of South Australia, 2012b) addressed two aspects of the clarifying forums: difficulties with the initial ones and the need for their continuation. They argued that the difficulties arose because of the newness of performance standards, inadequate exemplars, lack of presenters with subject expertise and attendees’ expectations. Participants in the current research had also identified poor exemplars and lack of presenter expertise and teacher expectations that the forums “would provide a detailed understanding of good assessment task design” (SACE Board of South Australia, 2012b, p. 82).

The Panel (SACE Board of South Australia, 2012b) reported that feedback on the 2012 clarifying forums had been “encouraging” (p. 82): 90% of completed exit surveys rated them as “useful or very useful” (p. 82). Although they acknowledged that only 35% of attendees completed the surveys, they appeared unaware that

similar proportions had reported positively on the 2011 forums (SACE Board of South Australia, 2011h).

They recommended not only the continuance of clarifying forums, but increasing their number and teacher “accessibility” (p. 6) to them and asked the SACE Board to reconsider increasing funding for teacher attendance. They suggested the forums “include a component of assessment task design” (p. 13).

Apparently the Panel saw an ongoing benefit in the clarifying forums, but failed to address the problem of funding greater teacher attendance.

### **8.6.6 Other topics mentioned in the evaluation and by participants**

The Report discussed a number of topics, not related to teacher learning, that had been mentioned by participants in the current research. Perceived biases against boys, students in government schools and students from working class backgrounds, were mentioned by one participant but other issues were addressed more frequently: moderation anonymity, the use of marks and statistical moderation.

The Report (SACE Board of South Australia, 2012b) confirmed the lack of anonymity of schools’ moderation samples that was reported previously by participants. The Report described this as concerning and recommended that steps be taken to ensure that moderators were not aware of the identity of the schools being moderated.

A few participants in the current research had lamented the lack of discrimination in SBA. John, for example, had lamented the “blunt nature” of performance standards as inappropriate for tertiary entrance selection and he would have preferred to be able to use a 100-point scale for the SBA. The First Year Evaluation (SACE Board of South Australia, 2012b) reported similar comments and proposed the SACE Board consider “the use of marks alongside grades” (p. 6), noting their use in other standards-based systems.

The report (SACE Board of South Australia, 2012b) commented that a number of submissions, “particularly those from teachers of mathematics and science” (p. 86) called for statistical moderation. This desire was prevalent in the current research with Physics teachers, who commonly described statistical moderation as “a very

good system” but only suggested it for the skills and applications tasks (SATs). The Panel appeared reluctant to support statistical moderation, arguing that “substantial changes ... [would be needed] to ensure a viable and valid procedure” (p. 86).

Among participants, preference for statistical moderation was based on two factors. First was the preference, noted in the literature (Duncan & Noonan, 2007; Klenowski & Wyatt-Smith, 2010a), for mathematics and science teachers to assess using numbers rather than descriptors. Second was participants’ observation of the similar natures of their SATs and the examination, arguing that the similarity made statistical moderation valid. It is noted, however, that the similarity observed in 2011 is susceptible to change. Although the participants restricted their SATs to tests, the syllabus (SACE Board of South Australia, 2011d) allows other possibilities, including multimedia products, oral presentations and interviews. Eventually, teachers may realise the flexibility inherent in the syllabus, so reducing the validity of statistically moderating SATs against the examination.

### **8.6.7 Summary of the first year evaluation**

The Evaluation Panel proclaimed a “key underpinning” of its report: an assessment system that values teacher judgement “must have as one of its primary goals the continued improvement” in teacher knowledge and understanding (SACE Board of South Australia, 2012b, p. 91). This had been evident in the Review that initiated the new SACE (Crafter et al., 2006), which argued that teachers would need to be provided with time, professional development and support materials “on a continuous basis” (p. 158). The First Year Evaluation argued for “continued improvement” (p. 91) in three areas: assessment task design, application of the performance standards and understanding of the reasons for moderation changes.

The recommendation in the initial SACE Review (Crafter et al., 2006) for “the establishment of an Institute for Educational Assessment” (p. 143) was repeated by the Evaluation Panel (SACE Board of South Australia, 2012b, p. 90) who noted that discussions had begun with universities and other stakeholders to this purpose. These discussions were reported in the SACE Board’s annual report of 2012 (SACE Board of South Australia, 2012a) and the proposal for its establishment was reported in 2013 (SACE Board of South Australia, 2013b). The Institute now has an active website (SACE Board of South Australia, 2014e).

It appeared that the Evaluation Panel (SACE Board of South Australia, 2012b) was more attuned to the wishes of the SACE Review (Crafter et al., 2006), particularly relating to ongoing teacher learning, than those charged with implementing the new SACE. It also appeared that the Evaluation Panel was aligned with the initial interpretation of the quality management cycle that underlay the current research—an expectation of annual extensive teacher involvement in the quality management processes.

## **8.7 Summary and recommendations**

The current research was undertaken into the potential educative value of the SACE quality management processes associated with SBA. The research found teacher learning had occurred in the area of assessment, but that teacher learning could be facilitated by designing the processes with their educative benefit as a priority rather than an incidental outcome. It was observed that some processes, assumed to be components in an annual cycle, were one-off introductory events. Consequently, recommendations are made to enhance the potential for teacher learning in three areas of SACE Board jurisdiction: induction programs, the quality management cycle and social moderation.

### **Induction programs**

- That induction programs be designed to encourage deeper understanding, with sufficient time for discussion to allow resolution of any dissonance.
- That there is greater subject teacher involvement in the planning of teacher development activities.
- That induction programs be implemented over a prolonged period and not restricted to introducing change.
- That induction programs be designed to facilitate and encourage teachers' experiential learning as they experiment and reflect on changes in understanding and practice.



### **Quality management cycle**

- That teacher learning be recognised as a major function of the quality management cycle and that the processes in the cycle be designed with this role in mind.
- That assessment tasks and student responses be used to stimulate online discussion among teachers.
- That website exemplars be supplemented and that teachers be engaged in the development of such materials.
- That ongoing quality management processes incorporate professional development in school-based assessment (e.g., the use of performance standards in task design, addressing bias and plagiarism, and improving the reliability of grading).

### **Social moderation**

- That social moderation continue for quality control of SACE school-based assessment.
- That all students in a subject in a school comprise one assessment group and that the school is responsible for intra-school moderation.
- That a greater number of teachers participate as moderators.
- That appointment of moderators be restricted to practising teachers.
- That well-designed professional development, possibly through the use of more exemplars and prolonged discussion, be provided to moderators to enhance the reliability of moderation decisions.
- That the nature of feedback provided in Queensland and New Zealand be explored with the aim of equipping SACE moderators to provide more informative feedback.
- That the moderator professional development used in Queensland and New Zealand be considered for equipping SACE moderators in providing more informative feedback.
- That moderators be afforded opportunities to develop their skills in writing school-specific feedback informing teachers of how their SBA may be improved.

- That the SACE Board cease payments to schools for releasing teachers for moderation and that the money be used for providing support moderators to work with subject teachers.
- That moderators be appointed for a period greater than one year.

It is believed that quality management processes are more likely to facilitate teacher learning if they are designed with teacher learning as a priority rather than as an incidental by-product.

The final chapter summarises the key findings of the current research. It describes the study's contributions to the literature and suggests areas of further research. The limitations of the study are discussed and reflections made on the research process.

## CHAPTER 9: CONCLUSION

The current research examined the quality management processes associated with school-based assessment (SBA) for their potential for teacher learning. Recent years have seen an increased use of SBA in high-stakes assessment in many systems (e.g., Briseid & Caillods, 2004; Crisp, 2013; Matters, 2006) and the research was undertaken in the context of the South Australian Certificate of Education (SACE) when changes were being introduced as a consequence of the SACE Review (Crafter et al., 2006). The Review had recommended the adoption of social moderation of SBA for its educative value for participating teachers. The current research differed from much previous research on social moderation (e.g., Adie et al., 2012; Klenowski et al., 2007; Wyatt-Smith & Castleton, 2005) by investigating a high-stakes environment and examining the possible educational value of other quality management processes, not only social moderation.

Four research questions addressed four aspects of the quality management processes: teacher involvement, the form of the processes, the content of the processes and evidence of teacher learning.

1. What is the extent of teacher involvement in the SACE Board's quality management processes and what teacher characteristics are associated with this involvement?
2. What evidence is there that the SACE Board's quality management processes possess the form or features of effective professional development that have been identified in the literature?
3. What evidence is there that the content of the SACE Board's quality management processes provides opportunities for teacher learning in the area of assessment?
4. What evidence is there of teacher learning that might be associated with involvement in the SACE Board's quality management processes?

The research focussed on teachers of SACE Physics, a subject with no history of social moderation. It was considered that these teachers might be more likely to

acquire, and demonstrate, learning in this area because it was the first time that SACE Physics assessment had incorporated social moderation. The data included SACE documents, observations of the quality management processes and a series of questionnaires and interviews with teachers over a twelve-month period, as well as two focus groups with SACE Board officers. The data were examined in the light of the literature on effective teacher and adult learning.

## 9.1 Key Findings

The data related to the first research question, *involvement*, revealed that, although some of the processes involved a majority of teachers, they were not part of an ongoing quality management cycle, but instead, constituted teacher induction to the assessment changes. Increased teacher involvement in the quality management processes had been viewed as desirable by the Review Panel (Crafter et al., 2006) for their potential educative value. The observed increase was minimal, meaning that few were afforded the opportunity of learning through involvement in social moderation (e.g., Adie et al., 2012; Connolly et al., 2012; Davidson, 2003; Ingvarson, 1990; Limbrick & Knight, 2005; L. Reid, 2007). Expectation of learning from the processes, particularly for the benefit of their students, was most commonly cited as the rationale for teacher involvement, supported by the literature on teacher learning (Guskey, 1986). The changes of 2011 stimulated online and inter-school collaborations, but these were short-lived. Intra-school collaborations were an extremely effective means of teacher learning, again reflecting the literature (e.g., Guskey, 2000; Hipkins & Robertson, 2011; Loucks-Horsley et al., 2003).

Data concerning the *form* of the quality management processes revealed that the teacher induction activities possessed some of the features associated with effective teacher learning, particularly relevance, coherence and the use of active learning (Desimone, 2009; Garet et al., 1999). It was considered, however, that they had failed to challenge teachers' pre-existing assumptions (Loucks-Horsley et al., 2003; Thompson & Zeuli, 1999; Timperley et al., 2007), were of limited duration (Desimone et al., 2002; Garet et al., 1999) and did not sufficiently consult teachers in the initial planning (Donavant, 2009; Loucks-Horsley et al., 2003; Zemke & Zemke, 1988). Participants who served as supervisors and moderators in the quality management cycle, reported learning from their experiences and these processes

satisfied the features described in the literature as necessary for effective teacher learning, particularly in their duration and discussion opportunities (Desimone et al., 2002; Garet et al., 1999; Hipkins & Robertson, 2011; Loucks-Horsley et al., 2003; Timperley et al., 2007). It was considered that engaging more teachers in ongoing learning about performance standards, particularly in task design and the grading of student work, would help to address the aspirations of the SACE Review (Crafter et al., 2006).

The *content* (the third research question) of both the induction program and the quality management cycle was relevant to teachers, with a particular focus on the use of performance standards. Participants, however, would have preferred more explicit instruction, particularly in assessment task design. Teachers were introduced to the changes expected of them, but without sufficient rationale for them. Although the exemplars were found to be useful, participants would have appreciated more. Participants were also very critical of the inadequate post-moderation feedback provided to teachers. While the Chief Assessor's report provided advice, because of its lack of specificity, teachers did not appreciate the relevance of comments to their own practices.

*Teacher learning* (Research Question 4) was evident in participants reporting changes in their assessment practices, particularly in investigations. Participants reported that their assessment changes commonly involved reflection on external sources, such as exemplar tasks and syllabuses, followed by iterative cycles of enactment and reflection. Participants demonstrated experiential learning (Kolb, 1984), reflective practice (Schon, 1987) and deep learning (Moon, 2000). Membership of a community of practice contributed to constructive reflection, particularly where there was more than one Physics teacher in a school. All participants acknowledged recognition of the importance of task design as something gained from the changes introduced in the SACE. A minority used performance standards in grading student work and most had used performance standards to check grades determined in other ways.

Although the current research was undertaken with a small sample of Physics teachers in the context of the SACE, there is enough evidence to indicate that a quality management cycle for SBA has the potential for teacher learning and that

learning would be facilitated if the processes were specifically designed for that purpose. This could be achieved by designing quality management processes based on evidence in the literature about (a) effective professional development and adult learning strategies and (b) research where social moderation has been found to contribute to teacher learning. Smaill (2013) argued that improvements in SBA should be facilitated by viewing teacher professional learning “as a goal, rather than a by-product, of social moderation” (p. 251). This argument is extended to all quality management processes conducted by the SACE Board.

## **9.2 Contribution to the literature**

Much literature on the educative value of social moderation has been in the primary and middle years of schooling (e.g., Adie et al., 2012; Klenowski et al., 2007; Limbrick & Knight, 2005; L. Reid, 2007; Wyatt-Smith & Castleton, 2005). The current research was conducted with teachers of students in their final year of secondary schooling, teachers who were very aware of the tertiary aspirations of most of their students. The tertiary entrance requirements of their students was an important issue for many of the participants and coloured their reception of the mandated changes. The research adds to the literature of moderation in a high-stakes environment (e.g., Hipkins & Hodgen, 2011; Hipkins & Robertson, 2011; Ingvarson, 1990; Maxwell, 2006; Smaill, 2013), reporting the learning described by moderators. Although moderators reported consistency in grading student work, the most common reports were of increased awareness of appropriate assessment tasks, increased confidence in one’s own assessment and improved presentation that might benefit students. In addition, by examining the potential for teacher learning in the other quality management processes, this research extends possibilities of teacher learning beyond that of social moderation alone.

The current research examined teacher learning through the lens of “micro-level episodes” (Evans, 2014), using the change sequence construct of the interconnected model of teacher learning (D. J. Clarke & Hollingsworth, 2002). Evidence supported the model of teacher learning through “mechanisms by which change in one domain is associated with change in another” (D. J. Clarke & Hollingsworth, 2002, p. 950). With the mandated changes as stimuli, and the quality management processes as resources, much participant change was not “done to teachers” (D. J. Clarke &

Hollingsworth, 2002, p. 948), but was better described as showing “teachers as learners” (D. J. Clarke & Hollingsworth, 2002, p. 949), learning through sequences of reflection and enactment: experiential learning (Kolb, 1984). In the research, mandated changes in assessment required teachers to change their practices. Subsequent changes in beliefs supported, at least in part, the argument that changes in teacher beliefs can follow changes in practice (Guskey, 1986).

### **9.3 Limitations of the study**

Limitations refer to weaknesses in research design (Punch, 2006), a number of which were inherent in the current research. First was the decision to restrict the research to Physics teachers. Although the selection of this subject appeared reasonable (see 3.3), awareness arose during the research that teachers’ preferred assessment paradigms are related to the subjects they teach (Duncan & Noonan, 2007; Klenowski & Wyatt-Smith, 2010a). Consequently, restricting the study to Physics teachers restricts the generalisability of participants’ beliefs and practices regarding performance standards to teachers of other subjects.

The small number of participants, thirteen, also limits the validity of any generalisations. Nevertheless, a range of opinions was collected from 39 questionnaires and 69 interviews with the same participants, who showed diversity in teaching experience, location and school sector, as well as the two focus groups of SACE Board officers. The need to recruit volunteers created a participant group that was not typical of SACE Physics teachers. While gender, career duration and Physics teaching experience might be considered typical, the high proportions of teachers from Independent schools and teachers with SACE Board assessment panel experience were atypical. Voluntary participation creates *selection-effects problems*, reported elsewhere (Donmoyer, 2012b), and it is probable that the current research suffered skewing due to the sample studied.

Data regarding changes in participant assessment practices were via self-reports, which may be inaccurate (J. P. Gall et al., 2005). Although a variety of sources was used, student data, which might improve reliability, were not used.

The collection of data during the first year of the new SACE introduced unforeseen complexities. The rationale for the timing was that teachers would be more aware of

changes in their assessment beliefs and practices in the first year of the new system. A confounding difficulty arose when the SACE Board processes changed during the research period. It had been assumed that planning and clarifying forums would be attended annually by most teachers. Instead, the revised system meant that most teachers experienced them only once, in their introduction to the new SACE.

When planning the verbalisation exercises, a decision was necessary concerning the tasks to be used. Three possibilities were to use a test, a practical investigation or an issues investigation. The decision to use a test proved unfortunate. Over the year, many participants changed the means of grading investigations but, because the examination was unchanged, participants continued to grade tests as previously. Consequently, evidence relating to participants' changed grading practices for investigations was restricted to self-reports and were not confirmed by the verbalisation exercises. The verbalisation exercises were not effective in revealing how participants made assessment decisions. Although others (Crisp, 2008b, 2010c; Greatorex & Suto, 2006) had allowed participants to practise verbalisation, practice was not allowed in the current research because it was suspected that practice might induce changes in behaviour. Previous verbalisation research has been with experienced markers whose practices were considered stable. It appeared that verbalisation is not an appropriate mechanism for studying changes in teacher marking practice. Alternative data collection methods, such as interviews or questionnaires (Crisp, 2013) might have been used. Although the verbalisation activities did not provide data on participants' grading processes, they did provide data on the grading outcomes. Participants' assigned grades, which were used in examining improved consistency, would not have been obtained with interviews or questionnaires.

The initial research design had planned to use questionnaires with many teachers. It was anticipated that changes might be evident in a number of variables derived from questionnaire data, such as participant confidence in the summative use of SBA. It was recognised that having only thirteen participants would be a serious impediment to obtaining reliable quantitative data from the questionnaires. Furthermore, there was not great variation between participants' responses to the Round 1 questionnaire items. When Rounds 1 and 3 data were compared, it was observed that participants



gave identical responses in both rounds with rare minor differences. Consequently, the quantitative data made a minimal contribution to the research. One possibility is that the representative nature of the sample may have contributed to the similarities in questionnaire responses. Another possibility is that the one-year duration for the research may have been insufficient for changes in beliefs to become evident. A third possibility lies in the quotation: “Perhaps, because the assessment is so closely tied with the rewards and penalties of their jobs, they could not afford to let loose of their original beliefs” (Yerrick et al., 1997, p. 155).

A number of years are required for teachers to fully implement changes in their teaching practices (Hipkins & Robertson, 2011; Loucks-Horsley et al., 2003; Wylie & Hodgen, 2010). This period was not available in the current research, which was restricted to twelve months. Even though, to some extent this may have been alleviated by examining the data for “micro-level episodes” (Evans, 2014) or change sequences (D. J. Clarke & Hollingsworth, 2002) in participants’ beliefs and practices, it is possible that participants’ practices and beliefs changed more after the first year than during it.

#### **9.4 Suggestions for further research**

This study was defined as a qualitative longitudinal study, but was limited by time constraints. An opportunity to interview the participants four years after the changes to SACE assessment might provide evidence of whether their assessment practices had continued to evolve. Had the changes of 2011 been discarded, maintained or extended?

The study found no evidence that differences in adoption of the changes were related to professional, school or personal characteristics of the participants. It is possible that differences in engagement with the change process might be linked to personal characteristics that were not investigated in the current research, such as “perceptions of their own self-worth in the form of doubts, resistance, disillusionment, and guilt” (van den Berg, 2002, p. 599). Study of such personal teacher characteristics and engagement with change might inform the implementation of other change processes.

The moderation feedback was criticised both by moderators writing it and by

teachers receiving it. Investigation into forms of moderation feedback, for example those provided in Queensland and New Zealand, might generate recommendations for improving such feedback. The literature on effective feedback (e.g. Hattie & Timperley, 2007) could also be applied to improving the quality of post-moderation feedback.

Much of the criticism of the unreliability of SBA appears to be based on the perceptions regarding the reliability of examinations (Hipkins, 2010b). While there has been some research into the reliability of moderation (Masters & McBryde, 1994), this is now quite dated. Although Queensland provides results of the quality review of its moderation (e.g., QSA, 2011), this appears to be unique. Comparisons between the reliability of examination markers and teachers appear desirable but, without access being provided by the appropriate authorities, they appear unlikely.

## **9.5 Reflections on the research process**

The data revealed that teachers changed their assessment practices and beliefs through a series of reflections and enactments. Their practices evolved over the research period and, possibly, continued to do so. The SACE assessment practices were also observed to evolve during the research period, sometimes causing participant frustration. Similar evolution was also evident in the research process. The final product differed in many ways from the initial proposal. Some changes were imposed, for example by the few teachers willing to participate. Other changes arose from literature written during the research, for example Evans's (2014) concept of "micro-level development", which stimulated the use of Clarke and Hollingsworth's interconnected model and change sequences (2002) for analysing changes described by participants. Other changes, for example in interview protocols, arose in reaction to events that occurred during the research, such as reflection on data and supervisor comments. The research revealed learning through evolutionary processes was common to teachers, institutions (such as the SACE Board) and researchers.

The influence of an observer on the observed is a concern (Patton, 2002) for researchers, such as myself, who believe that "social science inquiry should be objective" (R. B. Johnson & Onwuegbuzie, 2004, p. 14). I aimed to minimise my

influencing participants, attempting to capture the “external world” (Creswell, 2009, p. 11) without influencing it. Nevertheless, instances occurred where the research influenced participants. Elena, for example, in one interview referred to “whatever they’re called, you know all the different things” and hesitated, looking at me. I offered “the specific features”. Seven minutes later there was a similar exchange. After two more minutes, Elena correctly used the term “specific features” and continued to do so. It appeared that the interview contributed to her learning. It could be argued that I should not have responded to Elena’s nonverbal enquiry, but to ignore her could have jeopardised the relationship, negatively affecting the research in another way. In some instances, participants volunteered comments on the influence of the research. Adam, for example, observed “It’s been stimulating. It’s made me think.” In the final interview, when asked whether he had evaluated the changes replied “Probably only these discussions.”

## **9.6 Concluding Statement**

The data showed that teacher involvement in SACE activities increased greatly during 2011, with the changes to the SACE. It was also shown, however, that this was temporary, a consequence of the induction of teachers, rather than the ongoing increased involvement recommended in the SACE Review (Crafter et al., 2006). The data showed positive outcomes: teachers had changed their assessment practices, commonly stimulated by SACE Board documents, exemplars and discussions. The induction processes were generally effective, showing many of the features of effective professional development (e.g., Desimone, 2009; Garet et al., 1999; Timperley et al., 2007; Wei et al., 2009). Changes in teachers’ beliefs and practices were rarely immediate, instead evolving over time. As the First Year Evaluation (SACE Board of South Australia, 2012b) reported, teachers appreciated the opportunity “for professional development in the understanding and application of performance standards” (p. 102).

The participating teachers and SACE Board officers were intent on ensuring the changes to the SACE worked smoothly and were concerned for, in the case of the teachers, their students, or in the case of the officers, the teachers. The changes that teachers face are “highly complex” (Hargreaves et al., 2001, p. 13) and cannot be accomplished in a twelve-month period. It is important to note how teachers view

moderation (Hipkins, 2010a): a perception that it is “predominantly for accountability ... squanders the learning opportunity” (p. 31) inherent in social moderation discussions, which is what occurred here.

While participants described changes they had made and the change process, they appeared desperate for support, especially those who were the only Physics teacher in the school. Again, as the First Year Evaluation (SACE Board of South Australia, 2012b) said “there is still some way to go” (p. 102). For teachers of the SACE, the SACE Board offers the best opportunity for professional development in assessment and the quality management processes have the potential to satisfy this need. This is dependent upon whether the SACE Board perceives teacher learning as a priority of the quality management processes, as is the case in Queensland (QSA, 2009) and New Zealand (Ministry of Education, 2007) or whether teacher learning continues to be a by-product of the quality management processes.

# APPENDICES

## Appendix A: Teacher invitation

Dear Physics Teacher

This email invites you to take part in research into the potential for teachers to learn from involvement in the SACE Board quality assurances processes associated with school-based assessment of Stage 2 subjects.

I am undertaking research leading to the production of a thesis on this topic.

I would be most grateful if you would volunteer to assist in this project, by completing three online questionnaires over a 12-month period. It is estimated that 15-20 minutes will be required on each occasion.

In a second part of the research teachers are invited to participate in a series of eight 40-60 minute interviews over the 12-month period so that a deeper understanding can be gained. Willingness to participate in the online questionnaires does not oblige you to take part the interviews.

Be assured that any information provided will be treated in the strictest confidence and none of the participants will be individually identifiable in the resulting thesis, report or other publications. You are, of course, entirely free to discontinue your participation at any time or to decline to answer particular questions.

If you wish further information please contact me by email at Brent.Atherton@flinders.edu.au. Obtaining further information in no way obliges you to take any further part in the project.

If you are willing to participate you can log on to the first questionnaire at <http://edu.surveygizmo.com/s3/493322/Trial-v2>

Regards,

Brent Atherton

## **Appendix B: Officer invitation**

Dear SACE Board officer

This email invites you to take part in research into the potential for teachers to learn from involvement in the SACE Board quality assurances processes associated with school-based assessment of Stage 2 subjects.

I am undertaking research leading to the production of a thesis on this topic.

I would be most grateful if you would volunteer to assist in this project, by participating in two focus groups over a 12-month period. It is estimated that 40-60 minutes will be required on each occasion.

Be assured that any information provided will be treated in the strictest confidence and none of the participants will be individually identifiable in the resulting thesis, report or other publications. You are, of course, entirely free to discontinue your participation at any time or to decline to answer particular questions.

If you wish further information please contact me by email at Brent.Atherton@flinders.edu.au. Obtaining further information in no way obliges you to take any further part in the project.

If you are willing to participate please click on Reply and type YES before sending.

Regards,

Brent Atherton

Attachments: Letter of Introduction.doc and Research Information Sheet.doc.

---

## Appendix C: Teacher letter of introduction

Dear Physics Teacher

This letter is to introduce Mr Brent Atherton who is a Doctor of Philosophy student in the School of Education at Flinders University. He has a student card, which carries a photograph, as proof of identity.

He is undertaking research leading to the production of a thesis or other publications on the potential for teachers to learn from involvement in the SACE Board quality assurance processes associated with school-based assessment of Stage 2 subjects.

He will be contacting you by email and would be most grateful if you would volunteer to assist in this project, by completing three online questionnaires over a 12-month period. Approximately 15-20 minutes will be required on each occasion.

In a second part of the research teachers will be invited to participate in a series of 40-60 minute interviews so that a deeper understanding can be gained. Completion of the online questionnaires does not imply any obligation to undertake an interview.

Be assured that all information provided will be treated in the strictest confidence and none of the participants will be individually identifiable in the resulting thesis, report or other publications. You are, of course, entirely free to discontinue your participation at any time or to decline to answer particular questions.

Since he intends to make a recording of any interview, he will require your written consent to record the interview, to use the recording or a transcription in preparing the thesis, report or other publications, on condition that your name or identity is not revealed, and to make the recording available to other researchers on the same conditions. It may be necessary to make the recording available to secretarial assistants for transcription, in which case you may be assured that such persons will be advised of the requirement that your name or identity not be revealed and that the confidentiality of the material is respected and maintained. The transcript will be returned to you to allow confirmation of its contents prior to analysis.

Another part of the research involves SACE Board officers' observations of teacher involvement in the quality assurance processes.

Any enquiries you may have concerning this project should be directed to me at the address given above or by telephone on 8201 5459, or by email at [alan.russell@flinders.edu.au](mailto:alan.russell@flinders.edu.au).

Teachers who elect to participate in an interview have the right to decline to answer any question or to cease participation at any stage. In the event of a teacher becoming upset during an interview, the interview will be stopped. Should the teacher wish access to counselling, the following services are available:

DECS employees have access to the Employee Assistance Program through which a private company, Davidson Trahaire Corpsych Pty Ltd, delivers confidential counselling services to DECS staff. They can be contacted on 1300 360 364.

Employees in Catholic schools and most Independent schools have available the services of ACCESS-OCAR Services in Adelaide 08 (8231 9111 or 1300 66 77 00) and ACCESS in Whyalla at 08 8645 8233.

Thank you for your attention and assistance.

Yours sincerely

Emeritus Professor Alan Russell

School of Education

*This research project has been approved by the Flinders University Social and Behavioural Research Ethics Committee (Project Number 5192). For more information regarding ethical approval of the project the Executive Officer of the Committee can be contacted by telephone on 8201 3116, by fax on 8201 2035 or by email [human.researchethics@flinders.edu.au](mailto:human.researchethics@flinders.edu.au).*



## Appendix D: Officer letter of introduction

Dear SACE Board officer

This letter is to introduce Mr Brent Atherton who is a Doctor of Philosophy student in the School of Education at Flinders University. He has a student card, which carries a photograph, as proof of identity.

He is undertaking research leading to the production of a thesis or other publications on the potential for teachers to learn from involvement in the SACE Board quality assurances processes associated with school-based assessment of Stage 2 subjects.

He will be contacting you by email and would be most grateful if you would volunteer to assist in this project, by participating in two focus group sessions over a 12-month period. Approximately 50-60 minutes will be required on each occasion.

Be assured that all information provided will be treated in the strictest confidence and none of the participants will be individually identifiable in the resulting thesis, report or other publications. You are, of course, entirely free to discontinue your participation at any time or to decline to answer particular questions.

Since he intends to make a recording of each focus group session, he will require your written consent to record the session, to use the recording or a transcription in preparing the thesis, report or other publications, on condition that your name or identity is not revealed, and to make the recording available to other researchers on the same conditions. It may be necessary to make the recording available to secretarial assistants for transcription, in which case you may be assured that such persons will be advised of the requirement that your name or identity not be revealed and that the confidentiality of the material is respected and maintained. The transcripts will be returned to each participating officer to allow confirmation of its contents prior to analysis.

In a second part of the research teachers will be invited to complete three online questionnaires while a third part involves teachers participating in a series of 40-60 minute interviews so that a deeper understanding can be gained.

Any enquiries you may have concerning this project should be directed to me at the address given above or by telephone on 8201 5459, or by email at [alan.russell@flinders.edu.au](mailto:alan.russell@flinders.edu.au).

Thank you for your attention and assistance.

Yours sincerely

Emeritus Professor Alan Russell

School of Education

*This research project has been approved by the Flinders University Social and Behavioural Research Ethics Committee (Project Number 5192). For more information regarding ethical approval of the project the Executive Officer of the Committee can be contacted by telephone on 8201 3116, by fax on 8201 2035 or by email [human.researchethics@flinders.edu.au](mailto:human.researchethics@flinders.edu.au).*

---

## Appendix E: Research information sheet

### **Teacher learning through involvement in the quality management processes associated with social moderation**

This research is being undertaken by Brent Atherton, a full-time student in the School of Education who is studying for a Doctor of Philosophy degree. The Principal Supervisor is Emeritus Professor Alan Russell, a member of the staff of the School of Education at Flinders University.

The aim of the project is to investigate the potential for teachers to learn from involvement in the SACE Board quality assurance processes associated with school-based assessment of Stage 2 subjects. The processes to be considered for their potential impact on teacher learning include: planning forums, approval of learning and assessment plans, clarifying forums, moderation and feedback. Physics teachers are invited to participate. In the first part of the research, they are invited to complete three online questionnaires over a 12-month period. It is anticipated that each questionnaire will take 15-20 minutes to complete.

In the questionnaire, teachers will be invited to participate in a series of 40-60 minute interviews so that more detailed information can be obtained from a small number of teachers selected to provide a range of gender, experience, location, school size, and educational sector. All interviews will be conducted by the student-researcher, Brent Atherton. The questions and responses of each interview will be recorded with the participant's knowledge and permission. Afterwards a transcript will be made. It may be necessary to make the recording available to secretarial assistants for transcription, and such persons will be advised of the requirement that the name or identity of participants is not revealed and that the confidentiality of the material is respected and maintained. Each transcript will be returned to participants to allow the participant to edit it prior to analysis. Only participant-approved transcripts will be used in the project. Once the project is complete, each interview participant will be offered a copy of the report.

In a separate part of the research SACE Board officers involved in the quality assurance processes will also be invited to participate in two focus group sessions. The SACE Board officers will not have access to teachers' responses in either the questionnaires or the interviews.

All information provided will be treated in the strictest confidence and none of the participants will be individually identifiable in the resulting thesis, report or other publications. Participants are free to withdraw consent at any time or to decline to answer particular questions without needing to state a reason for doing so. If a teacher chooses to withdraw, all information already collected relating to that teacher will be destroyed and will not be used in the research.

The project is conducted under the auspices of the Social and Behavioural Research Ethics Committee of Flinders University which has ensured that the project complies with the Committee's ethical guidelines. Verification of this compliance can be obtained from the Executive Officer of the Social and Behavioural Research Ethics Committee, who may be contacted at the addresses shown above. Likewise, any other enquiries or complaints from participants or other interested parties should be through the Executive Officer.

Procedures for data collection, analysis, and storage shall ensure that it will not be possible for readers to identify individual teachers who have taken part in the research. Names and locations will be altered, and any identifying features, such as

teaching position, experience or gender will not be able to be linked to particular information.

Two backup copies will be made of each interview and focus group session and of each transcript. The recordings and transcripts will not be made available to other researchers.

The transcripts will be analysed to address the research questions which can be summarized as follows.

1. What is the extent of teachers' involvement in the SACE Board's quality management processes and what teacher characteristics are associated with this involvement?
2. What evidence is there that the SACE Board's quality management processes possess the features that have been shown, in the literature, to contribute to the effectiveness of professional development activities?
3. What evidence is there that the content of each of the SACE Board's quality management processes provides opportunities for teacher learning in the area of assessment?
4. What evidence is there that involvement with the SACE Board's quality management processes has contributed to teachers learning in the area of assessment?

Any enquiries you may have concerning this project should be directed to the Social and Behavioural Research Ethics Committee as described above, or to the supervisor or the researcher as described below.

Principal Supervisor

Researcher

Emeritus Professor Alan Russell

Brent Atherton

Phone: 8201 5459

Email: Brent.Atherton@flinders.edu.au

Email: alan.russell@flinders.edu.au

Member of academic staff

Student

School of Education

School of Education

Flinders University

Flinders University

*This research project has been approved by the Flinders University Social and Behavioural Research Ethics Committee. The Executive Officer of the Social and Behavioural Research Ethics Committee, can be contacted by telephone on 8201 3116, by fax on 8201 2035 or by email: [human.researchethics@flinders.edu.au](mailto:human.researchethics@flinders.edu.au).*

## Appendix F: Teacher consent form

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA

Social and Behavioural Research Ethics Committee

### CONSENT FORM FOR TEACHER PARTICIPATION IN RESEARCH

(by interview)

I .....  
 being over the age of 18 years hereby consent to participate as requested in the Letter of Introduction and Research Information Sheet for the research project on teacher learning arising from involvement in the quality assurance processes associated with school-based assessment.

1. I have read the information provided.
2. Details of procedures and any risks have been explained to my satisfaction.
3. I agree to my information and participation being recorded digitally on an MP3 player.
4. I am aware that I should retain a copy of the Information Sheet and Consent Form for future reference.
5. I understand that:
  - I may not directly benefit from taking part in this research.
  - I am free to withdraw from the project at any time and am free to decline to answer particular questions.
  - While the information gained in this study will be published as explained, I will not be identified, and individual information will remain confidential.
  - I am aware that it may be necessary to make the recording available to secretarial assistants for transcription, in which case such persons will be advised of the requirement that my name or identity not be revealed and that the confidentiality of the material is respected and maintained.
  - A transcript will be returned to me to allow me to review its contents and request changes prior to its analysis.
  - Neither the recording nor the transcript will be made available to other researchers.

**Participant's signature**.....**Date**.....

I certify that I have explained the study to the volunteer and consider that she/he understands what is involved and freely consents to participation.

**Researcher's name**.....

**Researcher's signature**.....**Date**.....

NB. Two signed copies should be obtained.

## Appendix G: Officer consent form

FLINDERS UNIVERSITY ADELAIDE • AUSTRALIA

Social and Behavioural Research Ethics Committee

### CONSENT FORM FOR PARTICIPATION IN RESEARCH

(by focus group)

I .....  
 being over the age of 18 years hereby consent to participate as requested in the Letter of Introduction and Research Information Sheet for the research project on teacher learning arising from involvement in the quality assurance processes associated with school-based assessment.

1. I have read the information provided.
2. Details of procedures and any risks have been explained to my satisfaction.
3. I agree to my information and participation being recorded digitally on an MP3 player.
4. I am aware that I should retain a copy of the Research Information Sheet and this Consent Form for future reference.
5. I understand that:
  - I may not directly benefit from taking part in this research.
  - I am free to withdraw from the project at any time and am free to decline to answer particular questions.
  - While the information gained in this study will be published as explained, I will not be identified, and individual information will remain confidential.
  - I acknowledge my obligation to maintain the anonymity of other members and the confidentiality of the discussion of the focus group session.
  - I am aware that it may be necessary to make the recording available to secretarial assistants for transcription, in which case such persons will be advised of the requirement that my name or identity not be revealed and that the confidentiality of the material is respected and maintained.
  - A transcript will be returned to me to allow me to review its contents and request changes prior to its analysis.
  - Neither the recording nor the transcript will be made available to other researchers.

**Participant's signature**.....**Date**.....

I certify that I have explained the study to the volunteer and consider that she/he understands what is involved and freely consents to participation.

**Researcher's name**.....

**Researcher's signature**.....**Date**.....

NB: Two signed copies should be obtained.

---

## Appendix H: Questionnaire 1 protocol

Thank you for undertaking this questionnaire. It is the first in a series of three that I would like you to complete over the next 12 months relating to the potential for teachers to learn from involvement in the SACE Board quality assurance processes.

All information you provide will remain confidential. No information will be released which would allow your identity to be known or inferred.

The questionnaire should take approximately 20 minutes to complete. There are five sets of questions: Assessment beliefs, Assessment practices, School expectations, SACE Board activities, and Personal information.

### Assessment beliefs

For each of the following items please select the response that best describes your personal beliefs about assessment: Strongly disagree, Disagree, Neutral, Agree, Strongly agree

- The prescribed curriculum is more important for planning lessons than how well students did in the last lesson
- The most useful assessment for student learning is assessment that is undertaken by the teacher rather than an external examination
- Assessment should help students to learn independently
- Students should be given guidance to assess their own work
- Students should be given guidance to assess other students' work
- Students benefit from opportunities to assess other students' work
- Assessment of student work should always be reported as marks
- Student effort is irrelevant when assessing their work
- It is important that students are told how they performed in relation to others in the class
- School-based assessment can assess more learning outcomes than external assessment
- School-based assessment is less rigorous than external assessment
- Students have too many opportunities to receive help with school-based assessment
- School-based assessment is open to manipulation by teachers
- There is a role conflict when a teacher is required to be an assessor
- I am not confident that other teachers make sure that the work they mark is the students' own work
- External assessment is the fairest form of assessment
- External assessment is objective
- A learning and assessment plan is merely bureaucratic; it serves no real educational purpose
- The performance standards are the foundation for designing assessment tasks
- Student work can be graded without the SACE performance standards
- Moderation is necessary to ensure that all teachers have followed the subject outline in their assessment
- Moderation is necessary to ensure that all teachers have followed their learning and assessment plan

- Moderation is necessary to ensure that students have undertaken comparable assessment tasks under similar conditions
- Moderation is necessary to ensure that students have not received undue help in assessment tasks
- Moderation is necessary to ensure that teachers have marked according to the performance standards
- I believe that some teachers will mark generously because moderation will not make small changes to teachers' grades
- I believe that my next learning and assessment plan will be approved without amendment
- I believe that my assessment tasks address all of the specific features of the subject outline
- I believe that my grading of student work is aligned with the standards of other teachers
- I believe that my interpretation of the performance standards is aligned with that of other teachers
- My assessment tasks allow students to demonstrate work at an A+ grade level

Please add any other comments you wish to make regarding your assessment beliefs.

### **Assessment practices**

For each of the following items please select the response that best describes your current assessment practices: Strongly disagree, Disagree, Neutral, Agree, Strongly agree.

- I adjust my teaching plan depending on how well students understood the work completed in class
- My assessment of students' work consists primarily of marks
- My assessment practices help students to learn independently
- I provide guidance to help students assess their own work
- I provide guidance to help students assess other students' work
- Students are given opportunities to assess other students' work
- I encourage students to view mistakes as valuable learning opportunities
- I use students' misconceptions shown in their work to adjust my teaching
- My assessment of student work includes extended written comments
- I use my knowledge of a student's effort when I assess their learning to determine their final mark
- When returning student work I tell students how well they have done in relation to others in the class
- I am able to design effective assessment tasks without referring to the assessment design criteria
- I find it hard to assess my students' work because my main role as a teacher is to help them develop educationally, emotionally and socially
- I make sure that my marking of each task is against the performance standards
- I make sure that I apply the performance standards in the same way to all students in my assessment
- When assessing I tend to be generous because I know that moderators will not be allowed to make small changes to students' grades

- 
- When assessing I tend to be a bit harsh because I do not want to risk my grades to be moderated downwards

Please add any other comments you wish to make regarding your assessment practices.

### **SCHOOL EXPECTATION**

For each of the following items please select the response that best describes your experience in your current school: Strongly disagree, Disagree, Neutral, Agree, Strongly agree.

- Discussion about school goals and means of achieving them is a regular part of our staff and faculty meetings
- My principal's values and philosophy of education are similar to my own
- There are a lot of irrelevant discussions at our faculty meetings
- I don't approve of the ways in which most of the other teachers in this school teach
- The principal of my school encourages teachers to talk with each other about instructional objectives
- Most teachers at my school have values and philosophies of education similar to my own
- At faculty meetings, we spend most of our time on the small stuff; we rarely get a chance to talk about the bigger issues in teaching and learning
- Teachers at this school share a high level of commitment to student learning
- I don't offer advice to others about their teaching unless I am asked for it
- My principal encourages me to try out new ideas
- Other teachers at this school come to me for help or advice when they need it
- New ideas presented at in-services are discussed afterwards by teachers in this school
- I can get good help or advice from other teachers at my school when I have a teaching problem
- I regularly share teaching ideas or materials with other teachers
- At this school, I have many opportunities to learn new things relevant to my teaching
- I give help and support to other teachers when they are having problems in their teaching
- I receive informal evaluations of my teaching performance from other teachers
- Other teachers encourage me to try out new ideas
- Other teachers at this school seek my advice about professional issues and problems
- When teachers are not doing a good job, our principal works with them to improve instruction
- My school encourages attendance at SACE Board teacher meetings
- My school encourages teachers to act as SACE Board markers or moderators
- Attendance at SACE Board teacher meetings is not encouraged in my school unless relief teachers are provided by the SACE Board
- Teachers in my school are not able to act as SACE Board moderators unless relief teachers are provided by the SACE Board



---

Please add any other comments you wish to make regarding your school's assessment expectations.

### **SACE BOARD ACTIVITIES**

#### **Planning forum**

If you attended a planning forum, for each of the following items please select the response that best describes your experience: Strongly disagree, Disagree, Neutral, Agree, Strongly agree

- The planning forum helped me to write an approved learning and assessment plan
- The planning forum helped me to design assessment tasks
- The planning forum helped me to use the performance standards in marking student work
- The planning forum helped me to apply the performance standards in a manner consistent with other teachers

If you were unable to attend a planning forum what was the reason?

#### **Learning and Assessment Plan approval**

Did you apply to be a panellist for Learning and Assessment Plan approval?  
Yes No

Was your Learning and Assessment Plan approved as submitted? Yes No

Did you check for and read any comments made about your Learning and Assessment Plan? Yes No

Did you contact a SACE Board officer for more detailed feedback about your Learning and Assessment Plan? Yes No

#### **Approval of learning and assessment plans**

If you approved learning and assessment plans, please select the response that best describes your experience: Strongly disagree, Disagree, Neutral, Agree, Strongly agree

- Approving other schools' learning and assessment plans will help me to write them in the future

#### **Feedback on learning and assessment plans**

Please select the response that best describes your experience of the feedback on your learning and assessment plan: Strongly disagree, Disagree, Neutral, Agree, Strongly agree

- Feedback on my learning and assessment plan will help me the next time I need to write one

Please add any other comments you wish to make regarding the planning forum and the learning and assessment plan approval process.

**PERSONAL INFORMATION**

Gender:      Female       Male

Years of teaching (including 2011):

Years of teaching Stage 2 Physics (including 2011):

Highest level of study in Physics:

Secondary      1 year tertiary      2 years tertiary      3 years tertiary      more than  
3 years tertiary

Where is your school located?

Metropolitan Adelaide       Regional South Australia

What sector does your school belong to?

Catholic       Government       Independent

The following questions relate to previous experience as a SACE Board Physics assessor (i.e., a setter, vetter, or marker) before 2011 when the new assessment processes were introduced.

How many years have you applied for a position as a SACE Board setter or vetter?

How many years have you served as a SACE Board setter or vetter?

How many years have you applied for a position as a SACE Board marker?

How many years have you served as a SACE Board marker?

=====

A unique identifier is needed for collation of responses to the three rounds of the questionnaire.

If you choose to give your name it will be known only to the researcher and will be kept separate from your other responses. The data collected will be used for statistical purposes and the identities of participating teachers will not be revealed.

You may prefer to remain anonymous by giving yourself an identification name or number that you will remember. For example, based on a birthdate of 3rd September 1981 your ID could be 03091981.

Name or identifier that you will remember \_\_\_\_\_

Are you willing to participate in a series of interviews over the next 12 months related to this research?

Yes  No

Your email address is needed to contact you if you are willing to be involved in a series of interviews over the next 12 months.

Email address: \_\_\_\_\_

## Appendix I: Questionnaire 2 protocol

Thank you for undertaking this questionnaire. It is the second in a series of three relating to the potential for teachers to learn from involvement in the SACE Board quality assurance processes.

All information you provide will remain confidential. No information will be released which would allow your identity to be known or inferred.

The questionnaire should take approximately 3 minutes to complete. There is one set of questions that relate to your experience of the SACE Board clarifying forums.

### SACE BOARD ACTIVITIES

#### Clarifying forum

Please complete the following items relating to your preparation for the clarifying forums held in Term 2 of 2011

Did you download the assessment exemplars before attending the clarifying forum?

Yes No

Did you grade the assessment exemplars before attending the clarifying forum?

Yes No

Did you submit your grades for the assessment exemplars to the SACE Board before attending the clarifying forum? Yes No

Did you submit any comments about the assessment exemplars to the SACE Board before attending the clarifying forum? Yes No

If you attended a clarifying forum, for each of the following items please select the response that best describes your experience: Strongly disagree, Disagree, Neutral, Agree, Strongly agree

- The clarifying forum helped me to write a learning and assessment plan
- The clarifying forum helped me to design assessment tasks
- The clarifying forum helped me to use the performance standards in marking student work
- The clarifying forum helped me to apply the performance standards in a manner consistent with other teachers

If you were unable to attend a clarifying forum what was the reason?

Please add any other comments you wish to make regarding the clarifying forum.

A unique identifier is needed for collation of responses to the three rounds of the questionnaire.

Name or identifier that you chose for the first questionnaire \_\_\_\_\_

If you did not complete the first round questionnaire, you may choose to give your name. It will be known only to the researcher and will be kept separate from your other responses. The data collected will be used for statistical purposes and the identities of participating teachers will not be revealed.

You may prefer to remain anonymous by giving yourself an identification name or number that you will remember. For example, based on a birthdate of 3rd September 1981 your ID could be 03091981.

Name or identifier that you will remember \_\_\_\_\_

---

## Appendix J: Questionnaire 3 protocol

Thank you for undertaking this questionnaire. It is the last in the series of three that you have completed over the last 12 months relating to the potential for teachers to learn from involvement in the SACE Board quality assurance processes.

All information you provide will remain confidential. No information will be released which would allow your identity to be known or inferred.

The questionnaire should take less than 15 minutes to complete. There are four sets of questions: Assessment beliefs, Assessment practices, your experiences of SACE Board activities, and changes in your assessment over the last two years.

### Assessment beliefs

For each of the following items please select the response that best describes your personal beliefs about assessment: Strongly disagree, Disagree, Neutral, Agree, Strongly agree

- The prescribed curriculum is more important for planning lessons than how well students did in the last lesson
- The most useful assessment for student learning is assessment that is undertaken by the teacher rather than an external examination
- Assessment should help students to learn independently
- Students should be given guidance to assess their own work
- Students should be given guidance to assess other students' work
- Students benefit from opportunities to assess other students' work
- Assessment of student work should always be reported as marks
- Student effort is irrelevant when assessing their work
- It is important that students are told how they performed in relation to others in the class
- School-based assessment can assess more learning outcomes than external assessment
- School-based assessment is less rigorous than external assessment
- Students have too many opportunities to receive help with school-based assessment
- School-based assessment is open to manipulation by teachers
- There is a role conflict when a teacher is required to be an assessor
- I am not confident that other teachers make sure that the work they mark is the students' own work
- External assessment is the fairest form of assessment
- External assessment is objective
- A learning and assessment plan is merely bureaucratic; it serves no real educational purpose
- The performance standards are the foundation for designing assessment tasks
- Student work can be graded without the SACE performance standards
- Moderation is necessary to ensure that all teachers have followed the subject outline in their assessment
- Moderation is necessary to ensure that all teachers have followed their learning and assessment plan

- Moderation is necessary to ensure that students have undertaken comparable assessment tasks under similar conditions
- Moderation is necessary to ensure that students have not received undue help in assessment tasks
- Moderation is necessary to ensure that teachers have marked according to the performance standards
- I believe that some teachers will mark generously because moderation will not make small changes to teachers' grades
- I believe that my next learning and assessment plan will be approved without amendment
- I believe that my assessment tasks address all of the specific features of the subject outline
- I believe that my grading of student work is aligned with the standards of other teachers
- I believe that my interpretation of the performance standards is aligned with that of other teachers
- My assessment tasks allow students to demonstrate work at an A+ grade level

Please add any other comments you wish to make regarding your assessment beliefs.

### **Assessment practices**

For each of the following items please select the response that best describes your current assessment practices: Strongly disagree, Disagree, Neutral, Agree, Strongly agree

- I adjust my teaching plan depending on how well students understood the work completed in class
- My assessment of students' work consists primarily of marks
- My assessment practices help students to learn independently
- I provide guidance to help students assess their own work
- I provide guidance to help students assess other students' work
- Students are given opportunities to assess other students' work
- I encourage students to view mistakes as valuable learning opportunities
- I use students' misconceptions shown in their work to adjust my teaching
- My assessment of student work includes extended written comments
- I use my knowledge of a student's effort when I assess their learning to determine their final mark
- When returning student work I tell students how well they have done in relation to others in the class
- I am able to design effective assessment tasks without referring to the assessment design criteria
- I find it hard to assess my students' work because my main role as a teacher is to help them develop educationally, emotionally and socially
- I make sure that my marking of each task is against the performance standards
- I make sure that I apply the performance standards in the same way to all students in my assessment
- When assessing I tend to be generous because I know that moderators will not be allowed to make small changes to students' grades

- When assessing I tend to be a bit harsh because I do not want to risk my grades to be moderated downwards

Please add any other comments you wish to make regarding your assessment practices.

## **SACE BOARD ACTIVITIES**

### **Part 1: Experience as a moderator**

Did you apply to be a moderator?    Yes    No

If you were a moderator, please complete the following items relating to your involvement in moderation that was undertaken at the end of 2011

- If you were a moderator, for each of the following items please select the response that best describes your current assessment practices: Strongly disagree, Disagree, Neutral, Agree, Strongly agree
- Serving as a moderator will help me to write an approved learning and assessment plan
- Serving as a moderator will help me to design assessment tasks
- Serving as a moderator will help me to use the performance standards in marking student work
- Serving as a moderator will help me to apply the performance standards in a manner consistent with other teachers

If you were a moderator, what other benefits did you gain from the experience?

Please add any other comments you wish to make regarding the end-of-year moderation.

### **Part 2: Feedback from moderation**

Were the top grade levels you submitted in Assessment Type 1: Investigations

Folio:

Moderated up    Unchanged    Moderated down

Were the middle grade levels you submitted in Assessment Type 1: Investigations

Folio:

Moderated up    Unchanged    Moderated down

Were the lowest grade levels you submitted in Assessment Type 1: Investigations

Folio:

Moderated up    Unchanged    Moderated down

Were the top grade levels you submitted in Assessment Type 2: Skills and Applications Tasks:

Moderated up    Unchanged    Moderated down

Were the middle grade levels you submitted in Assessment Type 2: Skills and Applications Tasks:

Moderated up    Unchanged    Moderated down

Were the lowest grade levels you submitted in Assessment Type 2: Skills and Applications Tasks:

Moderated up    Unchanged    Moderated down

For each of the following items please select the response that best describes the feedback that was sent to the school after moderation: Strongly disagree, Disagree, Neutral, Agree, Strongly agree

- The moderation feedback will help me to write an approved learning and assessment plan

- 
- The moderation feedback will help me to design assessment tasks
  - The moderation feedback will help me to use the performance standards in marking student work
  - The moderation feedback will help me to apply the performance standards in a manner consistent with other teachers

Did you contact a SACE Board officer for more detailed feedback about the moderation? Yes No

Please add any other comments you wish to make regarding the feedback provided after moderation.

### **Changes in assessment over the last two years**

Please select each of the following that applies to you over the last two years as far as assessment is concerned.

- I have questioned my beliefs about assessment.
- I have questioned the way I carry out assessment.
- I realized I no longer agreed with my previous beliefs.
- I realized I still agreed with my original beliefs.
- I realized that other teachers were also questioning their beliefs about assessment.
- I thought about carrying out my assessment in a different way.
- I felt uncomfortable with the previous form of assessment.
- I tried out new forms of assessment so that I would become more comfortable or confident with them.
- I tried out new forms of assessment so that I could select the best format.
- I tried to figure out a way to adopt new ways of assessing.
- I gathered information I needed to adopt new ways of assessing.
- I began to think about reactions and feedback about my new forms of assessment.
- I took action and adopted new ways of assessing.
- I changed my ways of assessing a number of times.

A unique identifier is needed for collation of responses to the three rounds of the questionnaire.

Name or identifier that you chose for the first two questionnaires \_\_\_\_\_

Thank you for completing this questionnaire. Your response is very valuable. I assure you that all information will be kept confidential and your anonymity maintained.

---

## Appendix K: Interview 1 protocol

### Introduction

I want to thank you for agreeing to share with me your understandings and experiences of assessment in Stage 2 Physics.

The questions are intended to be open-ended as I want to encourage you to talk about your experiences and thoughts. I plan to record this interview. Do you agree to my recording the interview?

I also ask you to sign two copies of the consent form so that each of us may retain a copy. I remind you that I intend to return a transcript to you so that you may amend it before I begin analysis.

If you are unsure about the meaning of any of the questions, please feel free to seek clarification. You also have the right to refuse to answer any individual question without needing to state your reason for doing so.

The interview covers three topics: your experience of the SACE Board quality assurance processes, how you learn about assessment, and your knowledge and beliefs about assessment.

### Experiences of quality management processes

The first set of questions relate to the planning forum and the approval process for the learning and assessment plans.

#### *Planning forum*

First, let's look at the planning forum; these were held back in the first half of 2010. I would like to look at the planning forum in two ways—attendance and your experience of the session itself. Let's look at attendance. In the questionnaire you said that ...

#### Lead question 1

What do you believe did the SACE Board hoped to achieve in the planning forums?

1. Could you describe your reasons for attending the planning forum?

OR

2. Did you expect that you might be missing out on anything by not attending the planning forum?
3. To what extent were you expected, encouraged, or required to attend the planning forum?
4. What did you expect to learn from attending the planning forum?
5. Were there any other benefits you expected from attending the planning forum?

#### Lead question 2

Now I ask you to recall your experiences of the planning forum—and how well they achieved the SACE Board's aims or what you hoped to gain from attendance.

6. What do you believe are the reasons behind the changes in assessment practices?
7. Did the planning forum provide any information on the reasons behind the changes in assessment practices?
8. Can you describe the activities that were used in the planning forum?



9. Could you describe opportunities that were given for questions to be addressed in the planning forum?
10. Could you describe opportunities that were given for discussion in the planning forum?
11. To what extent did the planning forum confirm or challenge your previous assessment beliefs or practices?
12. To what extent did you feel that teachers' points of view were listened to?
13. To what extent were you satisfied with the reasoning and principles that were used in the planning forum to explain such things as the changes in assessment practices?
14. To what extent were you satisfied with evidence that was given in the planning forum for conclusions and decisions made?
15. Was there anything that happened in the planning forum that caused you to do some more thinking, reading or discussing after the meeting?
16. Can you describe any change you have made, or may make, in your assessment because of the planning forum? What was it in the planning forum that could cause this change?
17. Have you experienced any sense of dissatisfaction or frustration in the changes you are required to make in Physics assessment this year?

### ***Learning and assessment plan approval***

Schools were required to submit a learning and assessment plan in all subjects either at the end of 2010 or early this year. These plans needed to be approved by a panel, mostly of teachers. I want to look at this process and the learning which might have been associated with it.

#### Lead question 3

In the questionnaire you said that the reason for the LAP was / was not merely bureaucratic.

What do you believe the SACE Board is trying to achieve with the LAP approval process and do you think it worked? What do you believe the SACE Board is trying to achieve with the approval process and do you think it worked?

18. In the questionnaire you stated that your school's LAP was ...  
Did you receive any information that will help you to complete your next LAP?
19. In the questionnaire you said that you did contact a SACE Board officer. What was the outcome of this?

OR

20. In the questionnaire you said that you did not contact a SACE Board officer. What was the reason for this?

#### Lead question 4

I would also like to ask about your involvement in the approval process. In the questionnaire you said that you applied / did not apply to be on the panel. What was the reason for this?—what were your reasons for (not) being involved?—what were your reasons for being involved and how would you describe the experience?

21. What did you expect to learn from being on the panel?
22. Were there any other benefits you expected from being on the panel?

- 
23. Can you describe any change you have made, or may make, in your assessment because of the panel involvement? What was it that could cause this change?

OR

24. Did you expect that you might be missing out on anything by not being on the panel?

### **Learning about assessment**

The second set of questions relate to your learning about assessment, both generally and from the SACE Board quality assurance cycle.

#### Lead question 5

I would like to ask you to look back over your time as a teacher and describe how you believe you have learned about assessment during that time.

25. In what ways has your assessment of students changed during your time as a teacher?
26. Were there any times at which change was sudden or has your assessment changed gradually over your years as a teacher?
27. What has caused you to change your assessment in the ways you describe?
28. Can you suggest what might be the best ways for you to continue learning about assessment?

### **Assessment knowledge and beliefs**

The final set of questions are planned to cover your knowledge and beliefs about assessment, particularly regarding the use of school-based assessment in Stage 2 Physics.

#### Lead question 6

I would like to ask you to comment on the use of school-based assessment in Stage 2 Physics and the moderation process that is being used this year.

29. In the questionnaire you said that  
External assessment is the fairest form of assessment  
School-based assessment is the most useful form of assessment  
Can you expand on this?
30. What are the advantages of including school-based assessment in the determination of students' final grades for the SACE?
31. What are the disadvantages of school-based assessment?
32. Do you believe that the extent of its use should be increased, decreased or remain as it is?
33. What are the advantages of using central moderation for adjusting school-based results in the determination of students' final grades for the SACE?
34. What are the disadvantages of central moderation?
35. Do you believe that central moderation should continue to be used, be modified in some way, or not used at all?
36. What reasons do you have for this belief?
37. In the questionnaire you said that you intend to be a little generous in the marking of student work. Could you expand on this?

38. What purposes are served by performance standards?
39. In the questionnaire you said that performance standards are useful in designing assessment tasks. Could you expand on this comment?
40. In the questionnaire you said that performance standards are useful in designing assessment tasks you can grade student work without the need for performance standards. Could you expand on this?
41. What are the disadvantages of using performance standards?
42. Do you believe that using performance standards should continue to be used, be modified in some way, or not used at all?
43. What changes would you like to see in the use of school-based assessment in the SACE?

---

## Appendix L: Interview 2 protocol

### Introduction

Thanks once again for agreeing to share with me your understandings and experiences of assessment in Stage 2 Physics. Do you agree to my recording the interview?

As with earlier interviews, the questions are intended to be open-ended as I want to encourage you to talk about your experiences and thoughts. If you are unsure about the meaning of any question, please feel free to seek clarification. You also have the right to refuse to answer any individual question or terminate the interview at any point without needing to state your reason for doing so.

The first part of the interview covers your experience of the SACE Board clarifying forum after which I would like to look at how you have gone about preparing materials for moderation and your expectations of that process.

First, can we look at the clarifying forum, your decision to attend and any preparation for the forum.

1. What do you believe the SACE Board was trying to achieve in the clarifying forum?
2. How well do you believe this was achieved?
3. What did you expect to learn from attending the clarifying forum?

OR

4. Did you expect that you might be missing out on anything by not attending the clarifying forum?
5. In the questionnaire you said that:
  - you did / did not download the exemplar materials that were made available online before the forum.
  - you graded / did not grade them
  - did / did not go online to register your assessment decisions.
  - did / did not go online to post any comments.

What were your reasons for this?

Now I ask you to recall your experiences of the clarifying forum—what happened, what were the outcomes for you?

In the clarifying forum it was made clear that decisions about standards had been made before the meeting.

6. Could you comment on the SACE Board using this process rather than allowing the decisions on standards to be made at the clarifying forum?
7. Who do you believe made the decisions about standards that were revealed at the clarifying forum? Were these the appropriate people to make the decisions and, if not, who should have made the decisions?
8. To what extent were the assessment decisions revealed at the clarifying forum aligned with the grades you had awarded?
9. To what extent were these decisions aligned with those of Physics teachers around the state?
10. To what extent were you satisfied with evidence that was given in the clarifying forum for conclusions and decisions made?

- 
11. To what extent did you find the assessment decisions confronting? How did you deal with this?
  12. Could you describe any issues which were raised in the clarifying forum?
  13. Could you describe opportunities that were given for questions and discussion in the clarifying forum?
  14. To what extent did you feel that teachers' points of view were listened to?
  15. To what extent did the clarifying forum confirm or challenge your previous assessment beliefs or practices?
  16. Was there anything that happened in the clarifying forum that caused you to do some more thinking, reading or discussing after the meeting?
  17. Can you describe anything you learned from the clarifying forum related to planning assessment, tasks design or use of the performance standards?
  18. Can you describe any change you have made, or may make, in your assessment because of the clarifying forum? What was it in the clarifying forum that could cause this change?

I would also like to ask about your involvement in the Clarifying Forum Support Panel.

19. Did you nominate to be on the support panel? Why was this?
20. What did you expect to learn from being on the panel?

OR

21. Did you expect that you might be missing out on anything by not being on the panel?
22. Can you describe any change you have made, or may make, in your assessment because of the panel involvement? What was it that could cause this change?

Can we now look at moderation—the processes you have gone through in preparation and your expectations.

I would like to look at each type of task—the Skills and Assessment Tasks, Practical Investigations, and Issues Investigation. To help me understand your descriptions, and to help remind you of your processes, I have copies of parts of the subject outline—the assessment design criteria, the specific features, and the performance standards.

23. First, could you describe how you arrived at the grade level you assigned each student for the Skills and Applications Tasks?
24. How did you use the specific features in designing your Skills and Applications Tasks?
25. How did you use the performance standards in arriving at the grade level for each student?

Second, could you describe how you arrived at the grade level you assigned each student for the Investigations Folio? For this, you may wish to describe separately the processes for the Practical Investigations from the processes used for the Issues Investigation.

26. How did you use the specific features in designing your Practical Investigations?
27. How did you use the specific features in designing your Issues Investigation?

- 
28. How did you use the performance standards in arriving at the grade level for each student across the two types of investigations?
  29. What materials are you going to provide to support the grades you have submitted?
- Third, can we look at the resources you have used in the assessment processes this year?
30. One resource that you were intended to use was the approved Learning and Assessment Plan that you submitted late last year or early this year. Could you describe the extent to which you have written addenda that describe variations on the approved plan.
  31. In what ways did the SACE Board exemplars stimulate, encourage or help you this year?
  32. In what ways did the SACE Board processes stimulate, encourage or help you in the processes you have used this year?
  33. What other resources (e.g., materials, people) have you used in developing the processes you have used this year?
  34. Finally, can we look at your expectations of the moderation process.
  35. Two statements have been made about the process by the SACE Board. The first of these is that the process involves moderation not re-marking. Have you heard this statement and what do you believe it means?
  36. The second statement is that moderation is about confirming the teacher's judgement. Have you heard this statement and what do you believe it means?
  37. During this year you have been required to make assessment judgements in task design and the grading of student work. How confident are you that your judgements will be confirmed?
  38. Have you already planned changes for next year? If so, what is the nature of the changes and what has led you to contemplate these changes?

---

## Appendix M: Interview 3 protocol

### Introduction

Thanks once again for agreeing to share with me your understandings and experiences of assessment in Stage 2 Physics. As in previous interviews the questions are intended to be open-ended as I want to encourage you to talk about your experiences and thoughts. If you are unsure about the meaning of any of the questions, please feel free to seek clarification.

You also have the right to refuse to answer any individual question without needing to state your reason for doing so. I plan to record this interview. Do you agree to my recording the interview?

The interview covers three topics: your experience of the SACE Board quality assurance processes; your knowledge and beliefs about assessment; and your learning about assessment as the changes in school-based assessment have been introduced over the last two years.

### Experiences of quality management processes

The first set of questions relate to the end-of-year moderation and the feedback that was provided.

#### *Moderation*

First, let's look at moderation.

#### Lead question 1

Could you describe your reasons for nominating or not nominating to be a moderator?

1. What led you to nominate or not nominate to be a moderator?
2. To what extent were you expected, encouraged, or required to nominate to be a moderator?
3. How willingly did you to make the decision to be or not be a moderator?
4. Did earlier experience on SACE Board panels influence your decision to nominate?
5. Did any knowledge about moderation influence your decision to (not) nominate to be a moderator?
6. Did any assumptions about moderation influence your decision to (not) nominate to be a moderator?
7. What did you expect to learn from being a moderator?
8. Were there any other benefits you expected from being a moderator?
9. Did you expect that you might be missing out on anything by not being a moderator?

#### Lead question 2

Now I ask you to you recall your experiences of being a moderator—what happened, what were the outcomes for you?

10. Can you describe the form of the activities that were used in training the panel?
11. Could you describe opportunities that were given for questions to be addressed in the panel training?

12. Could you describe opportunities that were given for discussion in the panel training?
13. How did the panel training confirm or challenge your previous assessment beliefs or practices?
14. Were you satisfied with the reasoning and principles that were used at moderation to explain the reasoning behind the process? Why?
15. Can you describe the process of moderation?
16. Were you satisfied with consistency of the moderation decisions made by different moderators? Why?
17. What happened at moderation to confirm or challenge your previous beliefs or practices about assessment?
18. What happened at moderation that caused you to do some more thinking, reading or discussing afterwards?
19. How might your experience as a moderator help you in your assessment this year?
20. How might your experience as a moderator help you to explain things to other teachers?
21. How has your experience as a moderator helped you to see connections between different parts of the assessment process?
22. How has your experience as a moderator helped you deal with any sense of dissatisfaction or frustration you had during last year?
23. What do you believe the SACE Board was trying to achieve in the training that was done at the beginning of the moderation?
24. How well do you believe this was achieved and why do you believe this?
25. What did you learn from being a moderator?
26. Did being a moderator reinforce anything you already knew? What did it reinforce? What happened to reinforce what you knew?
27. What did being a moderator clarify anything for you?
28. Can you describe any change you have made, or may make, in your assessment because your experience as a moderator? What was it in the moderation process that could cause this change?

### ***Feedback from moderation***

After moderation, feedback was sent to schools describing the outcome of the moderation process.

### **Lead question 3**

What do you believe the SACE Board is trying to achieve by providing the feedback and do you think it worked? How did it achieve this?

29. What was the outcome of moderation for your school? (Were your grade levels accepted or were changes made?)
30. What did you learn from the feedback that you received?
31. What acknowledgement did you receive that you had followed the procedures correctly?
32. What positive feedback did you receive about what you had submitted?



- 
33. What criticism did you receive about what you had submitted? Do you believe the criticism was justified and why?
  34. What advice did you receive about improvements you might make in the future?
  35. How detailed was the feedback?
  36. Were you able to understand the feedback?
  37. Are you aware that more detailed feedback may be available from the SACE Board officer in your subject?
  38. Did you follow up the feedback by getting in touch with anybody from the SACE Board?
  39. How accessible was the SACE Board officer?
  40. How approachable was the SACE Board officer?
  41. How helpful was the SACE Board officer?
  42. How satisfied were you with the response from the SACE Board officer and why?
  43. What was there in the feedback that reinforced something you already knew?
  44. What was there in the feedback that clarified something for you?
  45. What was there in the feedback that helped you to see links of which you had been unaware?
  46. What matters did the feedback raise which you believe you will need to spend time on?
  47. Was there any comment in the feedback that will lead you to change your practices in any way? If so, what was the comment and how do you think you might change your practices?
  48. The SACE Board has posted a Physics subject assessment report for 2011. Could you describe the use you have made of this report.

### **Assessment knowledge and beliefs**

The next set of questions are planned to cover your knowledge and beliefs about assessment, particularly regarding the use of school-based assessment in Stage 2 Physics.

#### Lead question 4

I would like to ask you to comment on the use of school-based assessment in Stage 2 Physics and the moderation process that was introduced last year.

49. What are the advantages of including school-based assessment in the determination of students' final grades for the SACE?
50. What are the disadvantages of school-based assessment?
51. Do you believe that the extent of its use should be increased, decreased or remain as it is?
52. What reasons do you have for this belief?
53. What are the advantages of using central moderation for adjusting school-based results in the determination of students' final grades for the SACE?
54. What are the disadvantages of central moderation?

- 
55. Do you believe that central moderation should continue to be used, be modified in some way, or not used at all?
  56. What reasons do you have for this belief?
  57. Could you describe how your beliefs about school-based assessment and moderation have changed over the past year?

#### Lead question 5

Can you now describe how you view the use of performance standards in school-based assessment in Stage 2 Physics.

58. Are performance standards necessary in writing assessment tasks? Why?
59. Are performance standards necessary in judging the quality of student work? Why?
60. What role do performance standards serve in judging the quality of student work?
61. What are the advantages of using performance standards?
62. What are the disadvantages of using performance standards?
63. Do you believe that using performance standards should continue to be used, be modified in some way, or not used at all?
64. What reasons do you have for this belief?
65. What changes would you like to see in the use of school-based assessment in the SACE?
66. What reasons do you give for suggesting these changes?

#### Lead question 6

Looking at your responses to the questionnaire, I noted that almost all of the responses were the same as last year or within one of last year's response. However, there were a few where your answer differed by two from last year's response. I would like to ask you about some of these.

#### **Learning about assessment**

The final set of questions relate to your learning about assessment over the last two years.

#### Lead question 7

I would like you to look back over the last two years and describe how your assessment has changed and what you believe you have learned about assessment during that time.

67. In what ways has your assessment of students changed during the last two years?
68. Were there any times at which you made a sudden change or did your assessment change gradually over the year?
69. What caused you to change your assessment in the ways you describe?
70. Do you feel a need to learn more about assessment or are you comfortable where you are? What makes you feel this way?
71. Can you describe the learning that you need to do?
72. Can you suggest what might be the best ways for you to continue learning about assessment?

Lead question 8

The questionnaire you completed recently asked you to make some selections about what has happened to you in the last two years as far as assessment is concerned. You ticked many of the statements as applying to you. It would appear that the time would have been fairly traumatic. How would you describe the changes that have occurred in you and how you carry out assessment?

73. Are there any events that stand out as being particularly significant over the last two years?
74. Can you describe any times at which you felt challenged to change what you were doing?
75. What actions did you take to increase your knowledge of the new requirements for assessment?
76. What actions did you take to help you understand the changes that were expected?
77. What actions did you take to help you see how the changes fitted in with what you already knew about assessment?
78. Are you aware of any point at which you evaluated the changes and made a decision that the new assessment was better or worse than what you were doing previously?

Lead question 9

Finally I would like you to respond to the recent announcement by the Minister of Education that there is to be a review into the new SACE.

79. If you were to make a submission to the review panel, what changes would you suggest?
80. What changes do you expect to see as a result of the review?

## Appendix N: Verbalisation protocol

### Verbalisation of assessment practices

I want to thank you for undertaking this assessment exercise for me and remind you that whatever happens is confidential and you will remain anonymous in any report of the session. The aim of the exercise is to gain information about the way that teachers undertake assessment activities.

To get information on ways in which you work when you are assessing, I want to observe you as you go about some tasks and I ask you to tell me what you are thinking to yourself, and saying to yourself, as you work on these tasks.

As far as is possible I would like you to tell me as much as you can of this thinking by talking aloud as you work on these problems.

The simplest way to do this is to **keep talking all the time** as you work just saying out loud what you are thinking.

**Try not to edit out** any of your thoughts, because these might turn out to have an important influence on the way you work.

Try to tell me everything you are thinking from the time you are given the first exercise until the end of the last one. Include as much **detail** as you can

Try to give me a **full report** of your thinking, of how you are going about the task—what you are thinking, what you will do first and so on. Change your mind as many times as you like, say if you think you need to do anything differently, and let me know when you have finished each exercise.

So try to **talk constantly**, all the time.

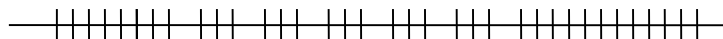
Just act as if you are in a room by yourself, and you are talking to yourself.

If you are silent for a time I will remind you to keep talking, so don't worry if I ask you to 'Keep talking', or 'Tell me what you are thinking.'

We begin with one exercise that has nothing to do with assessment to get you used to the idea of talking out loud as you are thinking. This is then followed by three assessment exercises.

We will do some practice in this thinking aloud now. Tell me all the details, however small. Even things that seem obvious. I'm interested in all your thinking.

Use Diagram



Imagine that all the little marks on this line are your thoughts as you are solving a problem. I want you to tell me as many of these as you can. In that walking task you mentioned this one (e.g., a line referring to going out the door) but did not mention this one (e.g., a line referring to going down a step).

You could try another exercise if you thought it necessary, or if the person understands the procedure go ahead.

Tell me all those details, however small.

I'm interested in all your thinking.

---

## Appendix O: Own task protocol

### Introduction

Three questions that I'm interested in are: What have you changed? Why did you change it the way you did? And what resources have you used in terms of other people, SACE Board information, and where have you got the information from for what you've done?

Follow-up questions

Where did that idea come from?

Did you bounce that idea off anyone?

Why did you made that change?

And where did you get the information about what you needed to do and the ideas about how to make your changes?

Have you used SACE Board materials at all to help you develop this?

How does it compare with last year's work?

So has this changed much from last year?

Is there anything you have done in this that you've done specifically because of the SACE performance standards?

Are you happy with the change?

What are the sorts of things that you have done to tasks this year to make them more acceptable?

Have the performance standards helped you?

And is that mainly from looking at that exemplar on the website?

---

## Appendix P: Round 1 focus group protocol

### Introduction

I want to thank you for agreeing to share with me your understandings and experiences of the quality assurance cycle for Stage 2 teachers. While my focus is on Physics teachers' learning, I am sure that observations made with teachers of other subjects will also provide information relevant to Physics teachers.

The questions are intended to be open-ended as I want to encourage you to talk about your experiences and thoughts. I plan to record this session. I would like each of you to acknowledge that you agree to my recording the session.

I also ask you to sign two copies of the consent form so that each of us may retain a copy. I remind you that I intend to return a transcript to each of you so that you may amend it before I begin analysis.

I remind you that each of you has an obligation to maintain the anonymity of other members and the confidentiality of the discussion of the focus group session.

If you are unsure about the meaning of any of the questions, please feel free to seek clarification. You also have the right to refuse to answer any individual question without needing to state your reason for doing so.

The focus group seeks your opinions and observations on three topics related to the SACE Board quality assurance cycle: teacher involvement, opportunities for teacher learning, and evidence of teacher learning.

### Planning forums

The first set of questions relate to teacher involvement in the planning forums and the approval process for the learning and assessment plans.

#### Lead question 1

I would like to hear your comments on teacher involvement in the planning forums—how many teachers attended, how willingly they attended, what happened if there was more than one teacher in a school who wanted to attend and anything else that comes to mind relating to attendance.

1. Were you aware of any teachers who wished to attend a meeting but were unable to do so? What were the reasons they were unable to attend?
2. Were you asked to provide information about any meeting to teachers who were unable to attend?
3. Were there instances of more than one teacher from a school attending the same meeting? How frequent were these?

#### Lead question 2

I would like to hear your comments on teacher learning that was intended in the planning forums.

4. What activities in the meetings did you think provided opportunities for teacher learning?
5. What was it about these activities that lead you to believe they provided the best opportunities?
6. What teacher learning did you expect to occur in these activities?

---

### Lead question 3

Could you comment on evidence you have that teachers have learned from the planning forums.

7. In what ways do teachers believe they have learned from the process?
8. Have you observed changes in teachers' practices as a result of the planning forums?
9. What evidence is there that the changes in teachers' practice are because of the planning forums?
10. Have teachers expressed a wish to have more opportunities similar to these?
11. Learning and Assessment Plan approval panel
12. Can we now look at those teachers who were panellists in the approval of learning and assessment plans.

### Lead question 4

Can you comment on the involvement of teachers in the panels—how many applied to be on the panels, how many were selected, what background did they have, how they were chosen?

13. Have any teachers talked about why they applied to be on the panel?
14. Have any teachers talked about schools encouraging or discouraging their involvement in the panel?

### Lead question 5

Before undertaking any approval, teachers went through a training process. Could you describe the learning that it was hoped that teachers would acquire from the panel training.

15. What activities were undertaken in the training that would develop teachers' understanding of the purpose of the learning and assessment plan?
16. What activities were undertaken in the training to ensure consistency in the approval process?

### Lead question 6

Could you describe the learning that teachers gained from being on the panel.

17. Have teachers spoken of benefits they received from the process?
18. Have any teachers talked of disadvantages that resulted from being on the panel?
19. Have teachers talked about changes they would like to see in the process?

### **Learning and Assessment Plan feedback**

The next topic is about teachers learning from the feedback they received to their learning and assessment plan.

### Lead question 7

What do you believe the SACE Board was trying to achieve with the feedback and do you think it worked?

20. What proportion of schools had their plans approved?
21. What proportion of schools had their plans approved but had suggestions made for improvements that could be made?

- 
22. How did the panellists respond to the feedback process—did they feel they would have liked to make more comments that were not allowed?
  23. How many schools have contacted you seeking an explanation of the feedback provided?
  24. How many schools have contacted you to complain about the feedback provided?
  25. How many schools have contacted you seeking additional feedback?
  26. How many schools have contacted you seeking help in addressing problems that had been identified in the feedback?
  27. How many schools have contacted you seeking an explanation of the feedback provided?

### **Clarifying forum support panel**

Can we now look at those teachers who were involved in the panels that provided support in the clarifying forums.

#### Lead question 8

Can you comment on the involvement of teachers in the panels—how many applied to be on the panels, how many were selected, what background did they have, how they were chosen?

28. Have any teachers talked about why they applied to be on the panel?
29. Have any teachers talked about schools encouraging or discouraging their involvement in the panel?

#### Lead question 9

Could you describe the learning that it was hoped that teachers would acquire from being on the panel.

30. What activities were undertaken in the training that would develop teachers' understanding of the purpose of the clarifying forum?
31. What activities were undertaken in the training for participation in the clarifying forums?

#### Lead question 10

Could you describe the learning that teachers gained from being on the support panel.

32. Have teachers spoken of benefits they received from being on the support panel?
33. Have any teachers talked of disadvantages that resulted from being on the panel?
34. Have teachers talked about changes they would like to see in the process?

### **Participation in clarifying forums**

The next topic is about teachers learning from participating in the clarifying forums.

#### Lead question 11

I would like to hear your comments on teacher involvement in the clarifying forums—how many teachers attended, how willingly they attended, what happened if there was more than one teacher in a school who wanted to attend and anything else that comes to mind relating to attendance.



35. What proportion of teachers attended a clarifying forum?
36. Were you aware of any teachers who wished to attend a meeting but were unable to do so? What were the reasons they were unable to attend?
37. Were you asked to provide information to teachers who were unable to attend?
38. Were there instances of more than one teacher from a school attending the same meeting? How frequent were these?

Lead question 12

I would like to hear your comments on teacher learning that was intended in the clarifying forums.

39. What activities in the meetings did you think provided opportunities for teacher learning?
40. What was it about these activities that lead you to believe they provided the best opportunities?
41. What teacher learning did you expect to occur in these activities?

Lead question 13

Could you comment on evidence you have that teachers have learned from the clarifying forums.

42. In what ways do teachers believe they have learned from the clarifying forums?
43. Have you observed changes in teachers' practices as a result of the quality management processes?
44. What evidence is there that teachers' practices are likely to change because of the clarifying forums?
45. Have teachers expressed a wish to have more opportunities similar to these?

---

## Appendix Q: Round 3 focus group protocol

### Introduction

Thanks again for agreeing to share with me your understandings and experiences of the quality assurance cycle for Stage 2 teachers. While my focus is on Physics teachers' learning, I am sure that observations made for teachers of other subjects will also provide information relevant to Physics teachers.

The questions are intended to be open-ended as I want to encourage you to talk about your experiences and thoughts. I plan to record this session. I would like each of you to acknowledge that you agree to my recording the session.

I remind you that each of you has an obligation to maintain the anonymity of other members and the confidentiality of the discussion of the focus group session.

If you are unsure about the meaning of any of the questions, please feel free to seek clarification. You also have the right to refuse to answer any individual question without needing to state your reason for doing so.

The focus group seeks your opinions and observations on three topics related to the SACE Board quality assurance cycle: teacher involvement, teacher learning from involvement as moderators, and teacher learning and from moderation feedback. Then I will ask you to look back over the last two years as far as teacher learning is concerned.

### Moderation

The first set of questions relate to teacher involvement in moderation that was held at the end of last year.

#### Lead question 1

I would like to hear your comments on teacher involvement in moderation—how easy was it to get moderators, how willingly they volunteered. Can you comment on the involvement of teachers in the panels—how many applied to be on the panels, how many were selected, what background did they have, how they were chosen?

1. Have any teachers talked about why they applied to be on the panel?
2. Have any teachers talked about schools encouraging or discouraging their involvement in the panel?
3. Were you aware of any teachers who wanted to be moderators but were unable to do so? What were the reasons they were unable to be moderators?
4. Were there instances of more than one teacher from a school being a moderator in the same subject? How frequently did this happen?
5. Were you aware of differences in involvement between the three sectors—Catholic, Government, Independent?
6. What steps are you aware of that were taken to increase involvement from any one of the sectors?
7. Were you aware of differences in involvement between country and metropolitan schools?
8. What steps are you aware of that were taken to increase involvement from teachers in country or metropolitan schools?

---

### Lead question 2

Before undertaking any moderation, teachers went through a training process. Could you describe the learning that it was hoped that teachers would acquire from the panel training.

9. What activities were undertaken in the training that would develop teachers' understanding of the moderation process?
10. What activities were undertaken in the training to ensure consistency in the moderation process?
11. What teacher learning did you expect to occur in this training?

### Lead question 3

Could you describe the learning that teachers gained from being on the panel and what evidence you have that teachers have learned from involvement as a moderator.

12. What benefits do teachers believe they have received from being a moderator?
13. In what ways do teachers believe they have learned from being a moderator?
14. Have any teachers talked of disadvantages that resulted from being on the panel?
15. Have teachers talked about changes they would like to see in the moderation process?
16. What opinions did moderators express about the moderation process and did these differ from the opinions expressed before moderation?
17. Are you expecting any changes in teachers' own assessment practices as a result of being a moderator?
18. What evidence is there that there will be changes in teachers' practice because of the moderation experience?
19. How many teachers have expressed a wish to be a moderator again?

### **Moderation feedback**

The next topic is about teachers learning from the feedback they received after moderation.

### Lead question 4

What do you believe the SACE Board was trying to achieve with the feedback and do you think it worked?

20. What proportion of schools had their grade levels changed in moderation?
21. How did the panellists respond to the feedback process—did they feel they would have liked to make more comments than were not allowed?
22. How many schools have contacted you seeking an explanation of the feedback provided?
23. How many schools have contacted you to complain about the feedback provided?
24. How many schools have contacted you seeking help in addressing problems that had been identified in the feedback?

25. Have you heard any opinions about whether teachers are more confident about the assessment processes now that one assessment cycle has been completed?

**Review of the last two years**

Lead question 5

I would like you to look back over the last two years and the changes that teachers have been required to make in their assessment. Could you comment on how well that has worked?

26. What changes have teachers made and what proportion of teachers have made the necessary changes?
27. How much resistance has there been from teachers to change their assessment? What reasons have they given for resisting the change? How has this resistance been expressed?
28. Have teachers acknowledged that their assessment is better than it was before the changes were made? In what ways do they believe their assessment has improved?
29. What resources have teachers said have been most helpful in making changes to their assessment practices?

Lead question 6

Finally I would like your response to the recent announcement by the Minister of Education that there is to be an evaluation of the new SACE.

30. From your experience with teachers over the last two years, what changes do you think they would you suggest to review panel?

## Appendix R: Questionnaire data analysis

### Confidence in formative

Table 41 lists the statements by James and Pedder (2006a) used to gather data on teachers' beliefs about SBA for creating student learning opportunities. The identifiers shown in brackets (e.g., A2) are those used by James and Pedder (2006a) for the statements in their questionnaires while the right hand column shows the labels used in the current research.

**Table 41: Items related to teachers' confidence in the formative use of SBA**

The prescribed curriculum is more important for planning lessons than how well students did in the last lesson (A2)	BF01
The most useful assessment for student learning is assessment that is undertaken by the teacher rather than an external examination (A8)	BF02
Assessment should help students to learn independently (A9)	BF03
Students should be given guidance to assess their own work (A13)	BF04
Students should be given guidance to assess other students' work (A19)	BF05
Students benefit from opportunities to assess other students' work (A29)	BF06
Assessment of student work should always be reported as marks (A12)	BF07
Student effort is irrelevant when assessing their work (A27)	BF08
It is important that students are told how they performed in relation to others in the class (A5)	BF09

It was considered that a teacher with confidence in the formative use of SBA would agree with the five items: BF02, BF03, BF04, BF05 and BF06. It is suggested that agreement with BF02 is self-evidently a person with confidence in the formative use of assessment. The other four items describe ways in which assessment may be used to enhance the learning process. James and Pedder (2006a) described BF03 as demonstrating belief that assessment is able to assist students "to develop independence in their learning" (p. 119) while they described BF04, BF05 and BF06 as *promoting learning autonomy* (p. 122).

The other four items (BF01, BF07, BF08 and BF09) see the purpose of assessment as measurement of learning outcomes and ranking of students, rather than as part of the learning process. BF01 emphasises the curriculum as the determinant of what is to be taught rather than student learning. The use of marks (BF07) is contrary to the provision of feedback that "focuses students on specific strategies for improvement" (Brookhart et al., 2010, p. 41) while the acknowledgement of student effort (BF08) contributes to the learning process as students see they are "capable of change and can accomplish it" (James & Pedder, 2006a, p. 117). BF09 describes the purpose of assessment as ranking rather than learning. Because these items do not demonstrate belief in assessment for formative purposes, they were reverse-coded.

SPSS analysis showed the effect on Cronbach's alpha of omitting items. Stepwise deletion of items to six (BF03, BF04, BF05, BF06, BF08R, and BF09R) generated a value for Cronbach's alpha of 0.502 and a mean inter-item correlation of 0.169.

### Confidence in summative

The items used for determining scores for *confidence in summative* were:

- 
- SBA can assess more learning outcomes than external assessment (BS01);
  - SBA is less rigorous than external assessment (BS02);
  - Students have too many opportunities to receive help with SBA (BS03);
  - SBA is open to manipulation by teachers (BS04);
  - There is a role conflict when a teacher is required to be an assessor (BS05).
  - I am not confident that other teachers make sure that the work they mark is the students' own work (BS06);
  - External assessment is the fairest form of assessment (BS07);
  - External assessment is objective (BS08);
  - A learning and assessment plan is merely bureaucratic; it serves no real educational purpose (BS09);
  - The performance standards are the foundation for designing assessment tasks (BS10);
  - Student work can be graded without the SACE performance standards (BS11).
  - I believe that some teachers will mark generously because moderation will not make small changes to teachers' grades (BS12).

BS01 and BS10 were considered favourable to the inclusion of SBA in high-stakes assessment. The other ten were considered to express an unfavourable opinion of the summative use of SBA and were reverse-coded.

Stepwise deletion identified nine items (BS03, BS04, BS05, BS06, BS07, BS09, BS10, BS11 and BS12) giving a Cronbach's alpha of 0.696 and a mean inter-item correlation of 0.229. The mean of these gave scores for *confidence in summative*. These were used in two ways: to analyse whether differences were associated with involvement in quality management processes and to see whether change had occurred over the period of research.

### **Formative practice**

The current research used eleven of James and Pedder's (2006a) statements to collect data on teachers' formative assessment practices. Table 42 shows the items used, with labels in brackets showing the labels used in James and Pedder's (2006a) instrument. Some statements were used in their original form, whereas others were reworded to make them more relevant to SACE Physics teachers.

**Table 42: Statements related to teachers' practices in the formative use of SBA**

I adjust my teaching plan depending on how well students understood the work completed in class (A1)	PF01
My assessment of students' work consists primarily of marks (A12)	PF02
My assessment practices help students to learn independently (A9)	PF03
I provide guidance to help students assess their own work (A13)	PF04
I provide guidance to help students assess other students' work (A19)	PF05
Students are given opportunities to assess other students' work (A29)	PF06
I encourage students to view mistakes as valuable learning opportunities (A16)	PF07
I use students' misconceptions shown in their work to adjust my teaching (A2)	PF08
My assessment of student work includes extended written comments (A22)	PF09
I use my knowledge of a student's effort when I assess their learning to determine their final mark (A27)	PF10
When returning student work I tell students how well they have done in relation to others in the class (A5)	PF11

Reverse-coding was used for PF02 and PF11. Cronbach's alpha for the eleven items was 0.278 while a mean inter-item correlation of 0.069. Omitting PF02 and PF11 gave an alpha value of 0.718 and a mean inter-item correlation of 0.241. The remaining items were used to generate *formative practice* scores.

### Summative practice

Table 43 shows the items that collected information on teachers' summative assessment practices.

**Table 43: Statements related to teachers' summative assessment practices**

I am able to design effective assessment tasks without referring to the assessment design criteria	PS01
I find it hard to assess my students' work because my main role as a teacher is to help them develop educationally, emotionally and socially	PS02
I make sure that my marking of each task is against the performance standards	PS03
I make sure that I apply the performance standards in the same way to all students in my assessment	PS04
When assessing I tend to be generous because I know that moderators will not be allowed to make small changes to students' grades	PS05
When assessing I tend to be a bit harsh because I do not want to risk my grades being moderated downwards	PS06

Analysis gave a value of 0.018 for Cronbach's alpha and a mean inter-item correlation of -0.002. It was expected that PS05 and PS06 would show a high negative correlation, suggesting that one of the items might be omitted. Although the value was negative, this combination was not the one with the most negative correlation. Because PS06 had negative inter-item correlations with all five of the

other items, it was reverse-coded. This generated a value of 0.382 for alpha and a mean inter-item correlation of 0.123. Omission of PS01 gave an alpha of 0.568 and a mean inter-item correlation of 0.215. The remaining five items were used to generate scores for *summative practice*.

### **School support**

Table 44 shows the school support items used from Rosenholtz (1989).



**School support**

**Table 44: School support items from Rosenholtz (1989)**

<b>Goal-setting</b>	
Discussion about school goals and means of achieving them is a regular part of our staff and faculty meetings	EG01
There are a lot of irrelevant discussions at our faculty meetings	EG02
At faculty meetings, we spend most of our time on the small stuff; we rarely get a chance to talk about the bigger issues in teaching and learning	EG03
The principal of my school encourages teachers to talk with each other about instructional objectives	EG04
<b>Shared goals</b>	
I don't approve of the ways in which most of the other teachers in this school teach	ES01
Most teachers at my school have values and philosophies of education similar to my own	ES02
Teachers at this school share a high level of commitment to student learning	ES03
My principal's values and education philosophy are similar to my own	ES04
<b>Collaboration</b>	
I don't offer advice to others about their teaching unless asked for it	EC01
Other teachers at this school come to me for advice when they need it	EC02
I can get good help or advice from other teachers at my school when I have a teaching problem	EC03
I give help and support to other teachers when they are having problems in their teaching	EC04
Other teachers at this school seek my advice about professional issues and problems	EC05
I regularly share teaching ideas or materials with other teachers	EC06
<b>Learning opportunities</b>	
New ideas presented at in-services are discussed afterwards by teachers in this school	EL01
At this school, I have many opportunities to learn new things relevant to my teaching	EL02
I receive informal evaluations of my teaching from other teachers	EL03
Other teachers encourage me to try out new ideas	EL04
My principal encourages me to try out new ideas	EL05
When teachers are not doing a good job, our principal works with them to improve instruction	EL06

A positive response (i.e., agreement) to Item EG02 was interpreted as a negative view of the school’s culture. Consequently it was reverse-coded, as were EG03, ES01, and EC01. Values were calculated for goal-setting as the mean of EG01–EG04. Values for shared goals were the mean of ES01–ES04. Values were calculated for collaboration as the mean of EC01–EC06 and learning opportunities values were the mean of EL01–EL06.

Table 45 shows the items written for collecting information relating to school support for SACE involvement.

**Table 45: Items relating to school support for SACE involvement**

My school encourages attendance at SACE Board teacher meetings	EB01
My school encourages teachers to act as SACE Board markers or moderators	EB02
Attendance at SACE Board teacher meetings is not encouraged in my school unless relief teachers are provided by the SACE Board	EB03
Teachers in my school are not able to act as SACE Board moderators unless relief teachers are provided by the SACE Board	EB04

Because the items had been taken from existing research (Rosenholtz, 1989) it was anticipated that correlations between the four variables *goal-setting*, *shared goals*, *collaboration* and *learning opportunities* might be high. However, since *school support for SACE* was generated for the current research, reliability analyses were undertaken for all items, with Table 46 showing the findings.

**Table 46: Reliability of school support data**

Derived variable	Contributing items	Cronbach’s alpha	Mean inter-item correlation
Goal-setting	EG01, EG02R, EG03R, EG04	0.808	0.533
Shared goals	ES01R, ES02, ES03, ES04	0.906	0.718
Collaboration	EC01R, EC02, EC03, EC04, EC05, EC06	0.754	0.442
Learning opportunities	EL01, EL02, EL03, EL04, EL05, EL06	0.942	0.771
School support for SACE	EB01, EB02, EB03R, EB04R	0.767	0.489
School support	Goal-setting, Shared goals, Collaboration, Learning opportunities, School support for SACE	0.901	0.681

These values are high enough to suggest a degree of redundancy. This was not pursued since school support formed such a small part of the data collected.

## Appendix S: Final coding for teacher learning

Table 47: Coding for Research Question 4

Level	Code
2	Active or passive learning
3	Active learning
3	Passive learning
2	Attitudes
3	Negative
4	Angry, annoyed
4	Concern over changes
4	Disillusioned
4	Dissatisfied with own effort
4	Disservice for students, worry
4	Frustration
4	Social moderation - lack of faith
4	Unsure, uncertain, insecure, confused, nervous, stressed
3	Positive
4	Confidence
4	Getting better
4	Hopeful
4	Open to change, compliant
2	Beliefs - LAP
2	Beliefs - Performance standards
3	Advantages - grading
4	Evidence for grades
4	Feedback for students
4	Help in marking
4	Inform students before task
3	Advantages - task design
4	Not needed for grading if used in task design
3	Disadvantages - grading
4	All in my head
4	Artificial for tests
4	Difficult to assess
4	Imprecise
4	Need explanations and exemplars
4	Negative effect on learning, education
4	Not understood by students
4	Open to interpretation
4	Time needed for administration
4	Time needed for teachers to learn
3	Disadvantages - task design
4	Cannot identify all that is important
4	Difficult to assess
4	Difficult to weight equally
4	Expertise needed
4	Open to interpretation
4	Time-consuming
4	Too many
4	Weighting needed
3	Purposes
4	Feedback to students
4	Identify what is important
4	Reasons explained
4	Rubrics skills-driven

4	Shift to rubrics
3	Similar grades to marks
3	Weighting
4	Should be equally weighted
2	Beliefs - SBA
3	Advantages
4	Can adapt to school situation
4	Feedback possible
4	Less stressful
4	More suitable for some tasks
4	More than one chance, can assess in greater depth
4	Some not good at exams
3	Disadvantages
4	Amount of work
4	Differences between schools
4	Less discrimination
4	Less rigorous
4	Whose work is it
2	Beliefs - Social moderation
3	Advantages
4	Good for non-test tasks
4	Good PD
4	Improves compliance
4	Shows what is happening in schools
4	Work is looked at, perceived to be fairer
3	Disadvantages
4	Administrative difficulties
4	Difficult to do, stressful for teachers
4	Educational disadvantages
4	Open to manipulation
4	Restricting school flexibility
4	Some teachers in the know
4	Subjective, lack of reliability, not valid
4	Upward moderation unlikely
3	Group moderation
3	Purposes
4	Adjusts for different schools
2	Deep learning
3	Reflection on change
4	Assessment changes
5	Assessment restricted
	Better assessment, better teaching
	Changed nature of assessment
	New way of thinking
4	Earlier years
4	Failure to question previous system, practices
4	Introduced too quickly, trial needed
4	Outcomes
5	Not worth the effort
5	Results the same
5	Standards changed
5	Washback
4	Professionalism
4	Purpose of education, assessment
4	Pushing students into the middle
4	Reflection on new SACE assumptions, beliefs and practices
4	Reflection on previous SACE assumptions, beliefs and practices
4	School requirements different from SACE

4	Subjectivity-objectivity
4	Weighting
3	Reflection on own practice
4	Failure to question own practice
4	Focus of SBA needs to change
4	Improved assessment
4	Weighting of specific features
3	Working with meaning
2	Learning from QM processes
3	Clarifying forum
3	Exemplars
4	Grading
4	Inadequate
4	Insufficient
4	LAP
4	Student use
4	Task design
4	Uncritical use
4	Used
3	Feedback from SACE Board
4	Chief Assessor's report
4	Insufficient
4	Learning expected
4	Nature of SACE Board feedback
3	Moderation
4	Different opinions, different ideas, different practices
4	Fair process
4	Good PD
4	Inside knowledge
4	Lack of reliability, flawed process
4	Looking good for moderation
4	Moderator training
5	Administration
5	Consistency achieved
5	More examples needed
4	Performance standards - improved knowledge
4	Task design
4	Upward moderation rare
3	Planning forum
3	SACE Board
4	Appreciation
4	Criticism
5	Changing during year
5	Lack of consistency
5	Lack of detail, specificity
5	Unwilling to assign specific features
4	Sympathy
2	Practice
3	Collaboration, cross-marking
3	Consistency
3	Feedback to students
4	Oral feedback
4	Performance standards
4	Written feedback
3	Grade cut-offs
4	15-point scale
4	85, 70, 55 cut-offs
3	Grading overall

4	Mathematical calculation of grades
5	Based on specific features
5	Based on tasks, total mark
5	Holistic review before moderation
5	Spreadsheet
5	Weighting of tasks
4	Non-mathematical review of grades
4	Weighting
3	Grading tasks
4	Based on marks for specific features
4	Based on performance standards
4	Based on total mark
4	Cut-offs decided after task
4	Evidence for grade
4	Marking harder
4	Performance standards completed after grade decided
4	Performance standards completed independent of marking
4	Professional judgement
3	LAP
3	Moderation preparation
4	Information provided
4	Looking good, helping moderators
3	Performance standards explained to students
3	Scheduling
3	Standards changed
3	Task design
4	Anonymity
4	Designed with moderation in mind, looking good
4	Guidance, information given
4	Higher order skills
4	Performance standards supplied with task sheet
4	Practical skills - more emphasis
4	Specific features
5	Assign specific features to task parts
6	Manipulated
6	Multiple specific features to a part
6	Weighting of specific features
5	Check coverage
4	Task simplification
3	Unchanged
2	Quality review
2	Teacher learning
3	Beliefs not changed
3	Better way
3	Challenged
3	Change sequence
3	Compliance
3	Factors and teacher learning
4	Affordances
5	Intra-school collaboration
5	In-school support
5	Inter-school collaboration
5	SACE meetings
5	Willing to adapt, desire to learn
4	Constraints
5	Alone
5	Change weary
5	Dissatisfaction with aspects of change

6	Research Project
6	SBA weighting
6	Subject choice
5	Lack of support
5	Previous new SACE experience
5	Questioning need for change
5	Resistance to change
5	Standards changed
5	Time demands
5	Unsure of requirements
3	Gradual change
3	Significant events
4	Clarifying forum
4	Inter-school collaboration
4	Justifying grade
4	Moderation
4	Staff meeting
3	Strategic approach
4	Playing the game
3	Support moderator
3	SACE terminology
2	Thesis discussion
3	Critique of research
3	Learning from research participation
3	Literature reference





## REFERENCES

- ACARA. (2012). The Australian Curriculum. Retrieved 30 December, 2012, from [http://www.acara.edu.au/curriculum/curriculum\\_design\\_and\\_development.html](http://www.acara.edu.au/curriculum/curriculum_design_and_development.html)
- ACER. (2012). Comparative Study of Selected Subjects for the Year 12 Certificate. Retrieved 30 December, 2012, from <http://www.acer.edu.au/research/projects/comparative-study-of-selected-subjects-for-the-year-12-certificate/>
- ACT Board of Senior Secondary Studies. (2012). Board of Senior Secondary Studies Policy and Procedures Manual 2012. Retrieved 15 December, 2012, from [http://www.bsss.act.edu.au/\\_\\_data/assets/pdf\\_file/0007/194821/01\\_PandPManual\\_2012\\_Amended\\_01-2012.pdf](http://www.bsss.act.edu.au/__data/assets/pdf_file/0007/194821/01_PandPManual_2012_Amended_01-2012.pdf)
- Adie, L. (2008). *Changing assessment practices: The case for online moderation*. Paper presented at the AARE 2008 International Education Research Conference. Retrieved from <http://www.aare.edu.au/08pap/adi08173.pdf>
- Adie, L. (2010). *Developing shared understandings of standards-based assessment: Online moderation practices across geographically diverse contexts*. Queensland University of Technology, Brisbane.
- Adie, L. (2011). An investigation into online moderation. *Assessment Matters*, 3, 5-27.
- Adie, L. (2013). The development of teacher assessment identity through participation in online moderation. *Assessment in Education: Principles, Policy & Practice*, 20(1), 96-106.
- Adie, L., Klenowski, V., & Wyatt-Smith, C. (2012). Towards an understanding of teacher judgement in the context of social moderation. *Educational Review*, 64(2), 223-240.
- Allal, L. (2013). Teachers' professional judgement in assessment: a cognitive act and a socially situated practice. *Assessment in Education: Principles, Policy & Practice*, 20(1), 20-34.
- Anderson, J., & Poole, M. E. (2001). *Assignment and thesis writing* (4th ed.). Milton, Qld.: John Wiley & Sons.
- Anderson, L. W., & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives* (Abridged ed.). New York: Longman.
- ARG. (2006). The role of teachers in the assessment of learning. Retrieved 5 August, 2008, from <http://www.assessment-reform-group.org/ASF%20booklet%20English.pdf>
- Atherton, B. (2009). *Teachers' understandings of, and attitudes towards, social moderation of school-based assessment*. Flinders University, Adelaide.
- Atherton, J. S. (2013). Learning and Teaching; SOLO taxonomy. Retrieved 1 November, 2010, from <http://www.learningandteaching.info/learning/solo.htm>
- Australian Institute for Teaching and School Leadership. (2011a). National Professional Standards for Principals. Retrieved 10 June, 2014, from [http://www.aitsl.edu.au/docs/default-source/default-document-library/australian\\_professional\\_standards\\_for\\_principals.pdf?sfvrsn=4](http://www.aitsl.edu.au/docs/default-source/default-document-library/australian_professional_standards_for_principals.pdf?sfvrsn=4)

- Australian Institute for Teaching and School Leadership. (2011b). National Professional Standards for Teachers. Retrieved 10 June, 2014, from [http://www.aitsl.edu.au/docs/default-source/default-document-library/aitsl\\_national\\_professional\\_standards\\_for\\_teachers](http://www.aitsl.edu.au/docs/default-source/default-document-library/aitsl_national_professional_standards_for_teachers)
- Baird, J. (2007). Alternative conceptions of comparability. In P. Newton, J. Baird, H. Goldstein, H. Patrick & P. Tymms (Eds.), (pp. 124-165). London: Qualifications and Curriculum Authority.
- Baird, J., Grooten, J., & Bell, J. F. (2004). What makes marking reliable? Experiments with UK examinations. *Assessment in Education*, 11(3), 331-348.
- Barrett, J. (2008). *The impact of the NCEA on teacher collegiality*. University of Waikato, Hamilton.
- Bennett, J., Tognolini, J., & Pickering, S. (2012). Establishing and applying performance standards for curriculum-based examinations. *Assessment in Education: Principles, Policy & Practice*, 19(3), 321-339.
- Bergman, M. M. (2010). On Concepts and Paradigms in Mixed Methods Research. *Journal of Mixed Methods Research*, 4(3), 171-175.
- Bergman, M. M. (2011a). The Good, the Bad, and the Ugly in Mixed Methods Research and Design. *Journal of Mixed Methods Research*, 5(4), 271-275.
- Bergman, M. M. (2011b). The Politics, Fashions, and Conventions of Research Methods. *Journal of Mixed Methods Research*, 5(2), 99-102.
- Biggs, J. B. (1999). What the Student Does: teaching for enhanced learning. *Higher Education Research & Development*, 18(1), 57-75.
- Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: the SOLO taxonomy (Structure of the Observed Learning Outcome)*. New York ; London: Academic Press.
- Biggs, J. B., Tang, C., & Society for Research into Higher Education. (2011). *Teaching for quality learning at university* (4th ed.). Maidenhead: McGraw-Hill/Society for Research into Higher Education/Open University Press.
- Bissaker, K. (2009). *The processes and outcomes of professional learning in an innovative school: the construction of an explanatory model*. Flinders University, Adelaide.
- Black, P. J. (1998). *Testing: friend or foe? The theory and practice of assessment and testing*. London: Falmer Press.
- Black, P. J., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). Assessment for Learning: putting it into practice. Retrieved 19 May 2010, from Open University Press: [http://www.gtce.org.uk/pdf/teachers/rft/afl\\_prac0904](http://www.gtce.org.uk/pdf/teachers/rft/afl_prac0904)
- Black, P. J., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: the classification of educational goals*. London: Longman Group.
- Board of Studies NSW. (2009, 19 June 2009). Performance Band descriptions Physics. Retrieved 21 January, 2011, from [http://www.boardofstudies.nsw.edu.au/syllabus\\_hsc/physics-pbd.html](http://www.boardofstudies.nsw.edu.au/syllabus_hsc/physics-pbd.html)
- Boereboom, J. (1999). *Assessment and Moderation of the Level 2 Physics Unit Standards on*

- 
- the National Qualifications Framework*. University of Canterbury, Christchurch.
- Bogdan, R., & Biklen, S. K. (1998). *Qualitative research for education: an introduction to theory and methods* (3rd ed.). Boston: Allyn and Bacon.
- Bogdan, R., & Biklen, S. K. (2007). *Qualitative research for education: an introduction to theories and methods* (5th ed.). Boston, Mass.: Pearson A & B.
- Booth, W. C., Colomb, G. G., & Williams, J. M. (2008). *The craft of research* (3rd ed.). Chicago: University of Chicago Press.
- Boyle, G. J. (1991). Does item homogeneity indicate internal consistency or item redundancy in psychometric scales? *Personality and Individual Differences*, 12(3), 291-294.
- Bransford, J., Brown, A. L., & Cocking, R. R. (Eds.). (1999). *How people learn: brain, mind, experience, and school*. Washington, D.C.: National Academy Press.
- Bransford, J., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: brain, mind, experience, and school*. Washington, D.C.: National Academy Press.
- Briggs, S. R., & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*, 54(1), 106-148.
- Briseid, O., & Caillods, F. (2004). Trends in secondary education in industrialized countries: Are they relevant for African countries? Retrieved 28 September, 2008, from <http://unesdoc.unesco.org/images/0013/001390/139018e.pdf>
- Broadfoot, P. M. (2005). Dark alleys and blind bends: testing the language of learning. *Language Testing*, 22(5), 123-141.
- Brock, S. E. (2010). Measuring the Importance of Precursor Steps to Transformative Learning. *Adult Education Quarterly*, 60(2), 122-142.
- Brookhart, S. M., Moss, C. M., & Long, B. A. (2010). Teacher inquiry into formative assessment practices in remedial reading classrooms. *Assessment in Education: Principles, Policy & Practice*, 17(1), 41-58.
- Brown, S. (1980). *What do they know? A review of criterion-referenced assessment*. Edinburgh: H.M. Stationery Office.
- Burns, R. (1785). To a Mouse, on Turning Her Up in Her Nest with the Plough.
- Burt, W. M., & Stapleton, L. M. (2010). Connotative Meanings of Student Performance Labels Used in Standard Setting. *Educational Measurement: Issues and Practice*, 29(4), 28-38.
- Charmaz, K. (2006). *Constructing grounded theory: a practical guide through qualitative analysis*. London ; Thousand Oaks, Calif.: Sage.
- Clarke, D. J., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, 18, 947-967.
- Clarke, D. J., & Peter, A. (1993). Modelling teacher change. In B. Atweh, C. Kanes, M. Carss & G. Booker (Eds.), *Contexts in Mathematics Education* (pp. 167-176). Kelvin Grove, Qld: MERGA.
- Clarke, E. (1987). Assessment in Queensland Secondary Schools: Two decades of change 1964-1983. Retrieved 25 April, 2014, from <http://education.qld.gov.au/library/docs/edhistory/assess/assess-01-14.pdf>
- Connolly, S., Klenowski, V., & Wyatt-Smith, C. M. (2012). Moderation and consistency of

- teacher judgement: teachers' views. *British Educational Research Journal*, 38(4), 593–614.
- Cooksey, R. W., Freebody, P., & Wyatt-Smith, C. (2007). Assessment as Judgment-in-Context: Analysing how teachers evaluate students' writing. *Educational Research and Evaluation*, 13(5), 401-434.
- Corbin, J. M., & Strauss, A. L. (2008). *Basics of qualitative research: techniques and procedures for developing grounded theory* (3rd ed.). Thousand Oaks, Calif.: Sage Publications Inc.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98-104.
- Cowie, B., Hipkins, R., Boyd, S., Bull, A., Keown, P., McGee, C. et al. (2009). *Curriculum implementation exploratory studies: Final report*. Wellington: Ministry of Education.
- Crafter, C., Crook, P., & Reid, A. (2006). *Success for all: Ministerial review of Senior Secondary Education in South Australia SACE Review Final Report*. Adelaide: Government of South Australia.
- Creswell, J. W. (2003). *Research Design Qualitative, Quantitative, and Mixed Methods Approaches* (2nd ed.). Thousand Oaks: Sage.
- Creswell, J. W. (2008). *Educational research: planning, conducting, and evaluating quantitative and qualitative research* (3rd ed.). Upper Saddle River, N.J.: Pearson/Merrill Prentice Hall.
- Creswell, J. W. (2009). *Research design: qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, Calif.: Sage Publications.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Los Angeles: Sage.
- Crick, R. D., Broadfoot, P., & Claxton, G. (2004). Developing an Effective Lifelong Learning Inventory: the ELLI Project. *Assessment in Education*, 11(3), 247-272.
- Crisp, V. (2008a). Exploring the nature of examiner thinking during the process of examination marking. *Cambridge Journal of Education*, 38(2), 247-264.
- Crisp, V. (2008b). The validity of using verbal protocol analysis to investigate the processes involved in examination marking. *Research in Education*, 79, 1-12.
- Crisp, V. (2010a). *The judgement processes involved in assessing GCSE coursework*. University of London, London.
- Crisp, V. (2010b). Judging the grade: exploring the judgement processes involved in examination grading decisions. *Evaluation & Research in Education*, 23(1), 19-35.
- Crisp, V. (2010c). Towards a model of the judgement process involved in examination marking. *Oxford Review of Education*, 36(1), 1-21.
- Crisp, V. (2013). Criteria, comparison and past experiences: how do teachers make judgements when marking coursework. *Assessment in Education: Principles, Policy & Practice*, 20(1), 127-144.
- Crisp, V., & Green, S. (2013). Teacher views on the effects of the change from coursework to controlled assessment in GCSEs. *Educational Research and Evaluation*, 19(8), 680-699.

- Crooks, T. J., Kane, M. T., & Cohen, A. S. (1996). Threats to the Valid Use of Assessments. *Assessment in Education: Principles, Policy & Practice*, 3(3), 265-285.
- Cross, J. (2007). *Informal learning: rediscovering the natural pathways that inspire innovation and performance*. San Francisco: Pfeiffer/Wiley.
- Cumming, J. J. (2010). Classroom Assessment in Policy Context (Australia). In B. McGaw, P. Peterson & E. Baker (Eds.), *The International Encyclopedia of Education* (Vol. 3, pp. 417-424). Oxford: Elsevier.
- Cumming, J. J., & Maxwell, G. S. (2004). Assessment in Australian schools: current practice and trends. *Assessment in Education*, 11(1), 89-108.
- Darling-Hammond, L., & McLaughlin, M. W. (1999). Investing in Teaching as a Learning Profession: Policy Problems and Prospects. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: handbook of policy and practice* (1st ed., pp. 376-411). San Francisco, Calif.: Jossey Bass Publishers.
- Daugherty, R. (1995). *National curriculum assessment: a review of policy, 1987-1994*. London: Falmer.
- Daugherty, R. (1997). Consistency in teachers' assessment: defining the problem, finding the answers. *British Journal of Curriculum and Assessment*, 8(1), 32-38.
- Davidson, J. (2003). All things in moderation: A whole school approach to authentic assessment. In H. Fehring (Ed.), *Literacy assessment: a collection of articles from the Australian Literacy Educators' Association* (pp. 81-83). Newark, Del.: International Reading Association.
- Davison, C. (2004). The contradictory culture of teacher-based assessment: ESL teacher assessment practices in Australian and Hong Kong secondary schools. *Language Testing*, 21(3), 305-334.
- Day, C., & Gu, Q. (2007). Variations in the conditions for teachers' professional learning and development: sustaining commitment and effectiveness over a career. *Oxford Review of Education*.
- De Vaus, D. A. (2002). *Surveys in social research* (5th ed.). St. Leonards, N.S.W.: Allen & Unwin Routledge.
- DEECD. (2007). Rubrics. Retrieved 21 January, 2011, from <http://www.education.vic.gov.au/studentlearning/assessment/preptoyear10/tools/rubrics.htm>
- DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, 17(4), 419-438.
- Denzin, N. K., & Lincoln, Y. S. (2005). *The SAGE handbook of qualitative research* (3rd ed.). Thousand Oaks: Sage Publications.
- Desimone, L. M. (2009). Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures. *Educational Researcher*, 38(3), 181-199.
- Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of Professional Development on Teachers' Instruction: Results from a Three-year Longitudinal Study. *Educational Evaluation and Policy Analysis*, 24(2), 81-112.

- 
- Díaz-Maggioli, G. (2004). *Teacher-centered professional development*. Alexandria, Va.: Association for Supervision and Curriculum Development.
- Donavant, B. W. (2009). The New, Modern Practice of Adult Education: Online Instruction in a Continuing Professional Education Setting. *Adult Education Quarterly*, 59(3), 227-245.
- Donmoyer, R. (2012a). Attributing Causality in Qualitative Research: Viable Option or Inappropriate Aspiration? An Introduction to a Collection of Papers. *Qualitative Inquiry*, 18(8), 651-654.
- Donmoyer, R. (2012b). Can Qualitative Researchers Answer Policymakers' What-Works Question? *Qualitative Inquiry*, 18(8), 662-673.
- Donmoyer, R., & Galloway, F. (2010). Reconsidering the Utility of Case Study Designs for Researching School Reform in a Neo-Scientific Era: Insights From a Multiyear, Mixed-Methods Study. *Educational Administration Quarterly*, 46(1), 3-30.
- Duncan, C. R., & Noonan, B. (2007). Factors Affecting Teachers' Grading and Assessment Practices. *The Alberta Journal of Educational Research*, 53(1), 1-21.
- Entwistle, N. (1988). Motivational Factors in Students' Approaches to Learning. In R. R. Schmeck (Ed.), *Learning strategies and learning styles* (pp. 21-51). New York: Plenum Press.
- Entwistle, N. (1998). Approaches to learning and forms of understanding. In B. Dart & G. Boulton-Lewis (Eds.), *Teaching and Learning in Higher Education* (pp. 72-101). Melbourne: Australian Council for Educational Research.
- Entwistle, N. (2000). *Promoting deep learning through teaching and assessment: conceptual frameworks and educational contexts*. Paper presented at the Teaching and Learning Research Programme.
- Entwistle, N., McCune, V., & Hounsell, J. (2002). *Approaches to Studying and Perceptions of University Teaching-Learning Environments: Concepts, Measures and Preliminary Findings*. Edinburgh: School of Education, University of Edinburgh.
- Eraut, M. (2007). Learning from other people in the workplace. *Oxford Review of Education*, 33(4), 403-422.
- Evans, L. (2014). Leadership for professional development and learning: enhancing our understanding of how teachers develop. *Cambridge Journal of Education*, 44(2), 179-198.
- Ezzy, D. (2002). *Qualitative analysis: practice and innovation*. Crows Nest, N.S.W.: Allen & Unwin.
- Fastier, J. M. L. (2007). *Teachers' perceptions regarding the implementation of the Level 1 Geography Achievement Standards within the New Zealand Senior Secondary School Context*. Christchurch College of Education, Christchurch.
- Fernandez, T. S. (2007). *From the Drawing Board into Schools: An Analysis of the Development and Implementation of a New Physics Curriculum in New Zealand Secondary Schools*. University of Waikato, Hamilton.
- Fernandez, T. S., Ritchie, G., & Barker, M. (2008). A sociocultural analysis of mandated curriculum change: the implementation of a new senior physics curriculum in New Zealand schools. *Journal of Curriculum Studies*, 40(2), 187-213.
-

- 
- Ferrari, J. (2007, 1 August 2007). No-exam Year 12 'inferior'. *The Australian*. Retrieved from <http://www.theaustralian.com.au/news/no-exam-year-12-inferior/story-e6frg6no-1111114082618>
- Fink, A. (2009). *How to conduct surveys: a step-by-step guide* (4th ed.). Thousand Oaks, Calif.: Sage.
- Foddy, W. H. (1993). *Constructing questions for interviews and questionnaires: theory and practice in social research*. Cambridge [England] ; Melbourne: Cambridge University Press.
- Fogarty, R., & Pete, B. (2004). *The adult learner: some things we know*. Thousand Oaks: Corwin.
- Fogarty, R., & Pete, B. (2010). Professional Learning 101: A Syllabus of Seven Protocols. *Phi Delta Kappan*, 91(4), 32-34.
- Fontana, A., & Frey, J. H. (2000). The interview: from structured questions to negotiated text. In N. K. Denzin & Y. S. Lincoln (Eds.), *The handbook of qualitative research* (2nd ed., pp. 645-672). Thousand Oaks, Calif.: Sage Publications.
- Forster, M. (2007). Best practice models for the development and use of performance standards. Retrieved 2 September, 2008, from [http://www.futuresace.sa.gov.au/a8\\_publish/modules/publish/content.asp?id=28580&navgrp=1430](http://www.futuresace.sa.gov.au/a8_publish/modules/publish/content.asp?id=28580&navgrp=1430)
- Fullan, M. (2007). *The new meaning of educational change* (4th ed.). New York: Teachers College Press.
- Fullan, M. (2008). *The six secrets of change: what the best leaders do to help their organizations survive and thrive* (1st ed.). San Francisco: Jossey-Bass.
- Fullan, M. (2011). Choosing the wrong drivers for whole system reform. Retrieved 24 December, 2012, from <http://www.michaelfullan.com/media/13436787590.html>
- Fullan, M., & Stiegelbauer, S. (1991). *The new meaning of educational change* (2nd ed.). London: Cassell.
- future SACE Office. (2008). Assessment Performance Standards and Moderation. Retrieved from <http://www.unisanet.unisa.edu.au/11788/sandm3/AssessmentPerformanceStan.pdf>
- Gall, J. P., Gall, M. D., & Borg, W. R. (2005). *Applying educational research: a practical guide* (5th ed.). New York: Longman.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: an introduction* (8th ed.). Boston: Pearson/Allyn & Bacon.
- Gardner, J. (Ed.). (2006). *Assessment and learning*. London: SAGE.
- Garet, M. S., Birman, B. F., Porter, A. C., Desimone, L. M., Herman, R., & Yoon, K. S. (1999). *Designing effective professional development: Lessons from the Eisenhower program*. Washington D.C.: U.S. Department of Education, Office of the Under Secretary, Planning and Evaluation Service
- Garet, M. S., Porter, A. C., Desimone, M., Birman, B. F., & Yoon, K. S. (2001). What Makes Professional Development Effective? Results From a National Sample of Teachers. *American Educational Research Journal*, 38(4), 915-945.
- Gibbs, G., & Simpson, C. (2004). Conditions Under Which Assessment Supports Students'

- Learning. *Learning and Teaching in Higher Education*, 1, 3-31.
- Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Katrien Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction*, 20, 304-315.
- Gipps, C. (1994a). *Beyond Testing*. London: The Falmer Press.
- Gipps, C. (1994b). *Quality Assurance in Teachers' Assessment*. Paper presented at the Annual Meeting American Educational Research Association.
- Gipps, C. (1995). What Do We Mean by Equity in Relation to Assessment? *Assessment in Education*, 2(3), 271-281.
- Gitlin, A., & Margonis, F. (1995). The Political Aspect of Reform: Teacher Resistance as Good Sense. *American Journal of Education*, 103(4), 377-405.
- Grammatikopoulos, V., Zachopoulou, E., Tsangaridou, N., Liukkonen, J., & Pickup, I. (2008). Applying a Mixed Method Design to Evaluate Training Seminars Within an Early Childhood Education Project. *Evaluation & Research in Education*, 21(1), 4-17.
- Greator, J., & Suto, W. M. I. (2006). *An empirical exploration of human judgement in the marking of school examinations*. Paper presented at the International Association for Educational Assessment Conference.
- Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of Coefficient Alpha as an Index of Test Unidimensionality. *Educational and Psychological Measurement*, 37, 827-838.
- Greene, J. C., Benjamin, L., & Goodyear, L. (2001). The Merits of Mixing Methods in Evaluation. *Evaluation*, 7(1), 25-44.
- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a Conceptual Framework for Mixed-Method Evaluation Designs. *Educational Evaluation and Policy Analysis*, 11(3), 255-274.
- Grix, J. (2004). *The foundations of research: a student's guide*. New York: Palgrave Macmillan.
- Guskey, T. R. (1986). Staff Development and the Process of Teacher Change. *Educational Researcher*, 15(5), 5-12.
- Guskey, T. R. (2000). *Evaluating professional development*. Thousand Oaks, Calif.: Corwin Press.
- Guskey, T. R. (2002). Professional Development and Teacher Change. *Teachers and Teaching: theory and practice*, 8(3), 381-391.
- Hammersley, M. (1997). Qualitative Data Archiving: Some Reflections on its Prospects and Problems. *Sociology*, 31(1), 131-142.
- Hargreaves, A., Earl, L., Moore, S., & Manning, S. (2001). *Learning to change: teaching beyond subjects and standards* (1st ed.). San Francisco: Jossey-Bass.
- Hargreaves, A., & Shirley, D. (2009). *The fourth way: the inspiring future for educational change*. Thousand Oaks, Calif.: Corwin Press.
- Harlen, W. (1994). *Concepts of quality in student assessment*. Paper presented at the American Educational Research Association Annual Meeting, New Orleans.
- Harlen, W. (2004a, March 2004). A systematic review of the evidence of reliability and validity of assessment by teachers used for summative purposes. *Research Evidence*



- in Education Library* Retrieved April 1, 2007, from [http://eppi.ioe.ac.uk/EPPIWeb/dl.aspx?Path=reel%5Creview\\_groups%5Cassessment%5Cass\\_rv3&FileName=ass\\_rv3](http://eppi.ioe.ac.uk/EPPIWeb/dl.aspx?Path=reel%5Creview_groups%5Cassessment%5Cass_rv3&FileName=ass_rv3)
- Harlen, W. (2004b). A systematic review of the evidence of the impact on students, teachers and the curriculum of the process of using assessment by teachers for summative purposes. Retrieved 17 May 2010, from EPPI-Centre, Social Science Research Unit, Institute of Education.: <http://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=6WO5QivP0Q4%3d&tabid=119&mid=925&language=en-US>
- Harlen, W. (2005). Teacher's summative practices and assessment for learning - tensions and synergies. *The Curriculum Journal*, 16(2), 207-223.
- Harlen, W. (2006). Role of Assessment in Developing Motivation for Learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 61-80). London: SAGE.
- Harlen, W. (2007). *Assessment of learning*. Los Angeles; London: SAGE.
- Harlen, W., & Deakin Crick, R. (2002). A systematic review of the impact of summative assessment and tests on students' motivation for learning. Retrieved 14 May 2010, from EPPI-Centre: <http://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=rOyQ%2ff4y3TI%3d&tabid=108&mid=1003&language=en-US>
- Hattie, J. (2003). *Teachers Make a Difference: What is the research evidence?* Paper presented at the Australian Council for Educational Research.
- Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. London; New York: Routledge.
- Hattie, J., & Timperley, H. S. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112.
- Hawley, W. D., & Valli, L. (1999). The Essentials of Effective Professional Development. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: handbook of policy and practice* (1st ed., pp. 127-150). San Francisco: Jossey-Bass Publishers.
- Heldsinger, S. A., & Humphry, S. M. (2013). Using calibrated exemplars in the teacher-assessment of writing: an empirical study. *Educational Research*, 55(3), 219-235.
- Hipkins, R. (2010a). The evolving NCEA. Retrieved 7 August, 2012, from <http://www.nzcer.org.nz/pdfs/evolving-ncea.pdf>
- Hipkins, R. (2010b). Learning through moderation: Minding our language. Retrieved 7 August, 2012, from <http://www.nzcer.org.nz/pdfs/hipkins-learning-through-moderation.pdf>
- Hipkins, R., Cowie, B., Boyd, S., Keown, P., & McGee, C. (2011). Curriculum Implementation Exploratory Studies 2: Final Report. Retrieved 6 August, 2013, from [http://www.educationcounts.govt.nz/\\_\\_data/assets/pdf\\_file/0005/91436/970\\_Implementation-Exploratory.pdf](http://www.educationcounts.govt.nz/__data/assets/pdf_file/0005/91436/970_Implementation-Exploratory.pdf)
- Hipkins, R., & Hodgen, E. (2011). *National Standards, moderation challenges and teacher learning*. Paper presented at the Symposium on Assessment and Learner Outcomes. Retrieved from [http://www.nzcer.org.nz/system/files/national-standards-moderation-challenges-teacher-learning\\_0.pdf](http://www.nzcer.org.nz/system/files/national-standards-moderation-challenges-teacher-learning_0.pdf)

- Hipkins, R., & Robertson, S. (2011). *Moderation and Teacher Learning: What can research tell us about their interrelationships?* Wellington: New Zealand Council for Educational Research.
- Hoekstra, A., & Korthagen, F. A. J. (2011). Teacher Learning in a Context of Educational Change: Informal Learning Versus Systematically Supported Learning. *Journal of Teacher Education*, 62(1), 76-92.
- Hollingsworth, H. (1999). *Teacher professional growth: A study of primary teachers involved in mathematics professional development*. Deakin University, Burwood, Victoria, Australia.
- Holstein, J. A., & Gubrium, J. F. (1997). Active Interviewing. In D. Silverman (Ed.), *Qualitative research: theory, method and practice* (pp. 113-129). London: Sage Publications.
- Holstein, J. A., & Gubrium, J. F. (2004). The active interview. In D. Silverman (Ed.), *Qualitative Research: Theory, Method and Practice* (2nd ed., pp. xiii, 378). London: Sage Publications.
- Ingvarson, L. (1990). *Enhancing professional skill and accountability in the assessment of student learning*. Paper presented at the Annual Meeting of the American Educational Research Association. Retrieved from <http://www.eric.ed.gov/PDFS/ED327558.pdf>
- James, M., & Pedder, D. (2006a). Beyond method: assessment and learning practices and values. *The Curriculum Journal*, 17(2), 109-138.
- James, M., & Pedder, D. (2006b). Professional Learning as a Condition for Assessment for Learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 27-44). London: SAGE.
- Jang, E. E., McDougall, D. E., Pollon, D., Herbert, M., & Russell, P. (2008). Integrative Mixed Methods Data Analytic Strategies in Research on School Success in Challenging Circumstances. *Journal of Mixed Methods Research* 2(3), 221-247.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, 33(7), 14-26.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research*, 1(2), 112-133.
- Johnson, S. (2011). A focus on teacher assessment reliability in GCSE and GCE. Retrieved 15 February 2013, from Office of Qualifications and Examinations Regulation: <http://www2.ofqual.gov.uk/files/reliability/11-03-16-A-focus-on-teacher-assessment-reliability-in-GCSE-and-GCE.pdf>
- Johnson, S. (2013). On the reliability of high-stakes teacher assessment. *Research Papers in Education*, 28(1), 91-105.
- Kagan, J., Segal, J., Havemann, E., Baucum, D., & Smith, C. D. (2004). *Kagan & Segal's psychology : an introduction* (9th ed.). Belmont, CA: Thomson/Wadsworth.
- Keeves, J. P. (1994). National examinations: design, procedures and reporting. Retrieved 5 February 2013, from UNESCO: International Institute for Educational Planning: <http://unesdoc.unesco.org/images/0009/000992/099209eo.pdf>
- Kilvert, P., & Mercurio, A. (2007). *Upper secondary level of education: who assesses?* Paper presented at the Education Research Conference.
- Klenowski, V., & Adie, L. (2009). Moderation as judgement practice: reconciling system

- level accountability and local level practice. *Curriculum Perspectives*, 29(1), 10-28.
- Klenowski, V., Adie, L., Gunn, S., Looney, A., Elwood, J., Wyatt-Smith, C. et al. (2007). *Moderation as Judgement Practice: Reconciling System Level Accountability and Local Level Practice*. Paper presented at the AARE 2007 International Education Research Conference *Research Impacts: Proving or Improving?* Retrieved from <http://eprints.qut.edu.au/17724/1/17724a.pdf>
- Klenowski, V., & Wyatt-Smith, C. (2010a). Standards-Driven Reform Years 1-10: Moderation an Optional Extra? *The Australian Educational Researcher*, 37(2), 21-39.
- Klenowski, V., & Wyatt-Smith, C. (2010b). Standards, teacher judgement and moderation in contexts of national curriculum and assessment reform. *Assessment Matters*, 2, 107-131.
- Knowles, M. S., Holton, E. F., & Swanson, R. A. (2005). *The Adult Learner* (6th ed.). Houston: Gulf Publishing Company.
- Kohn, A. (2007). Who's Cheating Whom? *Phi Delta Kappan*, 89(2), 89-97.
- Kolb, D. A. (1984). *Experiential learning: experience as the source of learning and development*. Englewood Cliffs, N.J.: Prentice-Hall.
- Kuijpers, J. M., Houtveen, A. A. M., & Wubbels, T. (2010). An integrated professional development model for effective teaching. *Teaching and Teacher Education*, 26, 1687-1694.
- Lave, J., & Wenger, E. (1991). *Situated learning: legitimate peripheral participation*. Cambridge England; New York: Cambridge University Press.
- Liamputtong, P., & Ezzy, D. (2005). *Qualitative research methods* (2nd ed.). South Melbourne: Oxford University Press.
- Limbrick, L., & Knight, N. (2005). Close reading of students' writing: What teachers learn about writing. *English Teaching: Practice and Critique*, 4(2), 5-22.
- Linn, R. L. (1993). Linking results of distinct assessments. *Applied Measurement in Education*, 6(1), 83-102.
- Little, J. W. (1999). Organizing Schools for Teacher Learning. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: handbook of policy and practice* (1st ed., pp. 233-262). San Francisco, Calif.: Jossey Bass Publishers.
- Long, C. (2006). *Realising the potential of school based assessment*. Paper presented at the Assessment in an Era of Rapid Change: Innovations and Best Practices, Singapore.
- Loucks-Horsley, S., Love, N., Stiles, K. E., Mundry, S., & Hewson, P. W. (2003). *Designing professional development for teachers of science and mathematics* (2nd ed.). Thousand Oaks, Calif.: Corwin Press.
- Lovett, S., & Gilmore, A. (2003). Teachers' Learning Journeys: The Quality Learning Circle as a Model of Professional Development. *School Effectiveness and School Improvement*, 14(2), 189-211.
- Luyten, H., & Dolkarb, D. (2010). School-based assessments in high-stakes examinations in Bhutan: a question of trust? Exploring inconsistencies between external exam scores, school-based assessments, detailed teacher ratings, and student self-ratings. *Educational Research and Evaluation*, 16(5), 421-435.
- Malone, L., Long, K., & De Lucchi, L. (2004). All Things in Moderation. *Science and*

- Children*, 41(5), 30-34.
- Marton, F., & Saljo, R. (1976). On Qualitative Differences in Learning — 1: Outcome and Process. *British Journal of Educational Psychology*, 46(1), 4-11.
- Marton, F., & Saljo, R. (2005). Approaches to Learning. In F. Marton, D. Hounsell & N. Entwistle (Eds.), *The Experience of Learning: Implications for teaching and studying in higher education* (3rd (Internet) edition ed., pp. 39-58). Edinburgh: University of Edinburgh, Centre for Teaching, Learning and Assessment.
- Masters, G. N., & McBryde, B. (1994). *An investigation of the comparability of teachers' assessment of student folios*. Brisbane: Queensland Tertiary Entrance Procedures Authority.
- Matters, G. (2006). *Statistical Moderation and Social Moderation around Australia*. Paper presented at the Assessment in an Era of Rapid Change: Innovations and Best Practices, Singapore.
- Maxwell, G. S. (2006). *Quality management of school-based assessments: Moderation of teacher judgments*. Paper presented at the Assessment in an Era of Rapid Change: Innovations and Best Practices, Singapore.
- Maxwell, G. S. (2009). Defining Standards for the 21st Century. In C. Wyatt-Smith & J. J. Cumming (Eds.), *Educational Assessment in the 21st Century: Connecting Theory and Practice* (pp. 269-293). Dordrecht Springer.
- Maxwell, G. S., & Cumming, J. J. (2011). Managing without public examinations: Successful and sustained curriculum and assessment reform in Queensland. In L. Yates, C. Collins & K. O'Connor (Eds.), *Australia's curriculum dilemmas: State perspectives and changing times* (pp. 202-222). Melbourne: Melbourne University Press.
- McGaw, B., Warry, R., & McBryde, B. (1975). Validation of Aptitude Measures for the Rescaling of School Assessment. *Education Research and Perspectives*, 2(2), 20-35.
- McTighe, J., & Emberger, M. (2008). Assessment as Professional Development. In L. B. Easton (Ed.), *Powerful designs for professional learning* (2nd ed., pp. 61-73). Oxford, OH: National Staff Development Council.
- Mercurio, A. (2003). *Questions as Answers: Understanding upper secondary selection and certification practices in South Australia, 1950 - 2000*. Wayville: Senior Secondary Assessment Board of South Australia.
- Mercurio, A. (2008, 10 September). *Re-imagining school-based assessment at the upper secondary education level*. Paper presented at the International Association for Educational Assessment, Cambridge.
- Merriam, S. B. (2001). Andragogy and self-directed learning: Pillars of adult learning theory. *New Directions for Adult and Continuing Education*, 2001(80), 3-13.
- Mezirow, J. (1994). Understanding Transformation Theory. *Adult Education Quarterly*, 44(4), 222-232.
- Miller, J., & Glassner, B. (1997). The 'Inside' and the 'Outside': Finding Realities in Interviews. In D. Silverman (Ed.), *Qualitative research: theory, method and practice* (pp. 99-112). London: Sage Publications.
- Mindich, D., & Lieberman, A. (2012). *Building a learning community: A tale of two schools*. Stanford: Stanford Center for Opportunity Policy in Education.

- Ministry of Education. (2007). *The New Zealand curriculum*. Wellington: Learning Media.
- Ministry of Education. (2011). *Ministry of Education position paper: Assessment [schooling sector]*. Retrieved from [http://www.minedu.govt.nz/~media/MinEdu/Files/TheMinistry/MOEAssessmentPositionPaper\\_October11.pdf](http://www.minedu.govt.nz/~media/MinEdu/Files/TheMinistry/MOEAssessmentPositionPaper_October11.pdf).
- Moon, J. A. (2000). *Reflection in learning & professional development: theory & practice*. London: Kogan Page.
- Morgan, D. L. (1997). *Focus groups as qualitative research* (2nd ed.). Thousand Oaks, Calif.: Sage Publications.
- National Quality Council. (2006). National Quality Council - Charter. Retrieved 19 February, 2010, from [http://www.dest.gov.au/NR/rdonlyres/F24807C9-92B8-4E33-9A89-8001B514920F/9340/nqc\\_charter1.pdf](http://www.dest.gov.au/NR/rdonlyres/F24807C9-92B8-4E33-9A89-8001B514920F/9340/nqc_charter1.pdf)
- National Quality Council. (2008). Code of Professional Practice for Validation and Moderation Available from [http://www.werc.vu.edu.au/publicAssets/pdf/Validation\\_and\\_Moderation\\_-\\_Code\\_of\\_Professional\\_Practice.pdf](http://www.werc.vu.edu.au/publicAssets/pdf/Validation_and_Moderation_-_Code_of_Professional_Practice.pdf)
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory*. New York: McGraw-Hill.
- OECD. (2005). Formative assessment: Improving learning in secondary classrooms. Retrieved 18 February, 2013, from <http://www.oecd.org/edu/ceri/35661078.pdf>
- OECD. (2009). Creating effective teaching and learning environments: First results from TALIS. Retrieved 17 August, 2010, from <http://www.oecd.org/dataoecd/17/51/43023606.pdf>
- Owens, L., Shute, L., & Slee, P. (2001). 'I'm in you're out ...'. *Psychology, Evolution & Gender*, 2(1), 19-46.
- Panizzon, D., Westwell, M., & Elliott, K. (2009). Updated Report: Who's teaching science in South Australia. Retrieved 22 May 2013, from Flinders Centre for Science Education in the 21st Century: [http://www.flinders.edu.au/science\\_engineering/science-21-files/briefings/whos\\_teaching\\_science.pdf](http://www.flinders.edu.au/science_engineering/science-21-files/briefings/whos_teaching_science.pdf)
- Patton, M. Q. (1987). *How to use qualitative methods in evaluation*. Newbury Park, Calif.: Sage Publications.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, Calif.: Sage Publications.
- Pedder, D., James, M., & MacBeath, J. (2005). How teachers value and practise professional learning. *Research Papers in Education*, 20(3), 209-243.
- Piedmont, R. L., & Hyland, M. E. (1993). Inter-Item Correlation Frequency Distribution Analysis: A Method for Evaluating Scale Dimensionality. *Educational and Psychological Measurement*, 53, 369-378.
- Pollard, A. (2002). *Reflective teaching: effective and evidence-informed professional practice*. London: Continuum.
- Popham, W. J. (1988). *Educational evaluation* (2nd ed.). Englewood Cliffs, N.J.: Prentice Hall.
- Porter, A. C., Garet, M. S., Desimone, L., Yoon, K. S., & Birman, B. F. (2000). *Does Professional Development Change Teaching Practice? Results from a Three-Year*

- 
- Study. (Report to the U.S. Department of Education, Office of the Under Secretary on Contract No. EA97001001 to the American Institutes for Research). Washington, DC: Pelavin Research Center.*
- Postholm, M. B. (2012). Teachers' professional development: a theoretical review. *Educational Research, 54*(4), 405-429.
- Punch, K. (2006). *Developing effective research proposals* (2nd ed.). London ; Thousand Oaks: SAGE.
- QCA. (2008). The Diploma glossary. Retrieved 4 October, 2014, from <http://dera.ioe.ac.uk/9411/>
- QSA. (2007a). Physics Senior Syllabus. Retrieved 21 January, 2011, from [http://www.qsa.qld.edu.au/downloads/senior/snr\\_physics\\_07\\_syll.pdf](http://www.qsa.qld.edu.au/downloads/senior/snr_physics_07_syll.pdf)
- QSA. (2007b). QSA districts. Retrieved 25 April, 2014, from [http://www.qsa.qld.edu.au/downloads/about/qa\\_districts\\_map.pdf](http://www.qsa.qld.edu.au/downloads/about/qa_districts_map.pdf)
- QSA. (2009). P-12 assessment policy. Retrieved 26 January, 2013, from Queensland Studies Authority: [http://www.qsa.qld.edu.au/downloads/approach/qa\\_assessment\\_policy.pdf](http://www.qsa.qld.edu.au/downloads/approach/qa_assessment_policy.pdf)
- QSA. (2010a). Moderation handbook for Authority subjects. Retrieved 31 May, 2012, from [http://www.qsa.qld.edu.au/downloads/senior/snr\\_moderation\\_handbook.pdf](http://www.qsa.qld.edu.au/downloads/senior/snr_moderation_handbook.pdf)
- QSA. (2010b). School-based assessment The Queensland system. Retrieved 30 October, 2011, from [http://www.qsa.qld.edu.au/downloads/approach/school-based\\_assess\\_qld\\_sys.pdf](http://www.qsa.qld.edu.au/downloads/approach/school-based_assess_qld_sys.pdf)
- QSA. (2010c). School-based assessment: The Queensland system. Retrieved 30 October, 2011, from [http://www.qsa.qld.edu.au/downloads/approach/school-based\\_assess\\_qld\\_sys.pdf](http://www.qsa.qld.edu.au/downloads/approach/school-based_assess_qld_sys.pdf)
- QSA. (2010d). Senior assessment general information. Retrieved 26 March, 2010, from <http://www.qsa.qld.edu.au/10-12/2130.html>
- QSA. (2010e). A-Z of Senior Moderation. Retrieved 27 April, 2010, from [http://www.qsa.qld.edu.au/downloads/10-12/snr\\_qa\\_mod\\_a-z.pdf](http://www.qsa.qld.edu.au/downloads/10-12/snr_qa_mod_a-z.pdf)
- QSA. (2011). *Random sampling project: 2011 Report on random sampling of assessment in Authority subjects*. Brisbane: Queensland Studies Authority.
- QSA. (2013). Membership of state or district review panel: Application form. Retrieved 5 May, 2014, from <https://www.qsa.qld.edu.au/1426.html>
- Rapley, T. (2004). Interviews. In C. e. a. Seale (Ed.), *Qualitative research practice* (pp. 15-33). London: Sage.
- Rawlins, P., Brandon, J., Chapman, J., Leach, L., Neutze, G., Scott, A. et al. (2005). Standards-Based Assessment in the Senior Secondary School: A Review of the Literature. *New Zealand Journal of Teachers' Work, 2*(2), 107-115.
- Reid, A. (2009, 2-4 October 2009). *Is this a revolution?: A critical analysis of the Rudd government's national education agenda*. Paper presented at the 2009 ACSA National Biennial Conference; Curriculum: a national conversation, Canberra, ACT.
- Reid, L. (2007). Teachers talking about writing assessment: valuable professional learning? *Improving Schools, 10*, 132-149.
-

- Rennie, D. L. (2000). Grounded Theory Methodology as Methodical Hermeneutics: Reconciling Realism and Relativism. *Theory Psychology, 10*(4), 481-502.
- Richardson, V., & Placier, P. (2001). Teacher change. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 905-947). Washington, D.C.: American Educational Research Association.
- Rolph, L., & Jordan, P. (2010). *Externally moderated school-based assessment in Queensland: lessons for implementing an Australian curriculum*. Paper presented at the Assessment for the future generations. Retrieved from <http://selectscore.com/fullpaper/511.pdf>
- Rosenholtz, S. J. (1989). *Teachers' workplace: the social organization of schools*. New York: Longman.
- Rosenholtz, S. J., & Simpson, C. (1990). Workplace Conditions and the Rise and Fall of Teachers' Commitment. *Sociology of Education, 63*(4), 241-257.
- Russell, A., & Russell, G. (1994). Coparenting Early School-Age Children: An Examination of Mother-Father Interdependence Within Families. *Developmental Psychology, 30*(5), 757-770.
- Ryan, M. T., Irwin, J. A., Bannon, F. J., Mulholland, C. W., & Baird, A. W. (2004). Observations of veterinary medicine students' approaches to study in preclinical years. *Journal of Veterinary Medicine Education, 31*(3), 242-254.
- SACE Board of South Australia. (2009). *Stage 1 and Stage 2 Physics Curriculum Statements 2009*. Wayville: SACE Board of South Australia.
- SACE Board of South Australia. (2010a). *2010 (Term 1) Stage 2 subject workshops: Workshop evaluation survey feedback*. Wayville: SACE Board of South Australia.
- SACE Board of South Australia. (2010b). *2010 (Term 2) Stage 2 subject workshops: Workshop evaluation survey feedback*. Wayville: SACE Board of South Australia.
- SACE Board of South Australia. (2010c). *2010 Stage 2 Workshops: Workshop evaluation survey, SACE Board of South Australia*. Wayville.
- SACE Board of South Australia. (2010d). *Annual Report 31 December 2010*. Wayville: SACE Board of South Australia.
- SACE Board of South Australia. (2010e). *Booklet 1 Physics Stage 2 Workshop*. Wayville: SACE Board of South Australia.
- SACE Board of South Australia. (2010f, February 2010). Quality Assurance in the SACE. *SACE Board of SA Policy Circulars* Retrieved 13 July, 2010, from [http://www.sace.sa.edu.au/c/document\\_library/get\\_file?p\\_1\\_id=244552&folderId=77775&name=DLFE-8307.pdf](http://www.sace.sa.edu.au/c/document_library/get_file?p_1_id=244552&folderId=77775&name=DLFE-8307.pdf)
- SACE Board of South Australia (Producer). (2010g) *SACE Assessment and Reporting: Guidelines for Teachers*.
- SACE Board of South Australia. (2010h). *SACE Assessment Responsibility*. Retrieved February 27, 2011, from <http://www.sace.sa.edu.au/>
- SACE Board of South Australia. (2010i). *SACE Moderation Procedures*. Retrieved 19 February, 2010, from [http://www.sace.sa.edu.au/c/document\\_library/get\\_file?uuid=f1de8040-d67c-4777-9a07-97425489bed5&groupId=10336](http://www.sace.sa.edu.au/c/document_library/get_file?uuid=f1de8040-d67c-4777-9a07-97425489bed5&groupId=10336)

- 
- SACE Board of South Australia. (2010j). *SACE Operations Manual 2010*. Wayville: SACE Board of South Australia.
- SACE Board of South Australia. (2010k). Stage 2 Assessment and Quality Assurance in 2011. Retrieved 24 November, 2010, from [http://www.sace.sa.edu.au/c/document\\_library/get\\_file?p\\_l\\_id=244552&folderId=77775&name=DLFE-23302.pdf](http://www.sace.sa.edu.au/c/document_library/get_file?p_l_id=244552&folderId=77775&name=DLFE-23302.pdf)
- SACE Board of South Australia. (2010l). *Stage 2 Physics Curriculum Statement 2010*. Wayville: SACE Board of South Australia.
- SACE Board of South Australia. (2010m). Stage 2 Quality Assurance & Assessment Panels in 2011. Retrieved 15 November, 2010, from [http://www.sace.sa.edu.au/c/document\\_library/get\\_file?uuid=2fecb410-98ec-44fa-ba23-ee2a5a641f14&groupId=10336](http://www.sace.sa.edu.au/c/document_library/get_file?uuid=2fecb410-98ec-44fa-ba23-ee2a5a641f14&groupId=10336)
- SACE Board of South Australia. (2010n). Stage 2 Quality Assurance and Assessment Panel Opportunities in 2011. Retrieved 15 November, 2010, from <http://www.sace.sa.edu.au/about/office-of-the-sace-board/stage-2-qa-and-assessment-panels-2011>
- SACE Board of South Australia. (2011a). 2011 Physics Enrolment Counts. Wayville, South Australia: SACE Board of South Australia.
- SACE Board of South Australia. (2011b). *Annual Report 31 December 2011*. Wayville: SACE Board of South Australia.
- SACE Board of South Australia. (2011c). Online forum Physics Retrieved 3 March, 2011, from <http://www.sace.sa.edu.au/group/physics>
- SACE Board of South Australia. (2011d). Physics 2011 Subject Outline Stage 1 and Stage 2. Retrieved 17 December, 2010, from [http://www.sace.sa.edu.au/c/document\\_library/get\\_file?p\\_l\\_id=10718&folderId=61824&name=DLFE-20613.pdf](http://www.sace.sa.edu.au/c/document_library/get_file?p_l_id=10718&folderId=61824&name=DLFE-20613.pdf)
- SACE Board of South Australia. (2011e, February 2010). Quality assurance in the SACE. *SACE Board of SA Policy Circulars* Retrieved 13 July, 2010, from [http://www.sace.sa.edu.au/c/document\\_library/get\\_file?p\\_l\\_id=244552&folderId=77775&name=DLFE-8307.pdf](http://www.sace.sa.edu.au/c/document_library/get_file?p_l_id=244552&folderId=77775&name=DLFE-8307.pdf)
- SACE Board of South Australia. (2011f). *SACE Operations Manual 2011*. Wayville: SACE Board of South Australia.
- SACE Board of South Australia. (2011g). Sciences 2011 Learning Area Manual. Retrieved 22 January, 2011, from [http://www.sace.sa.edu.au/c/document\\_library/get\\_file?p\\_l\\_id=10718&folderId=288048&name=DLFE-27601.pdf](http://www.sace.sa.edu.au/c/document_library/get_file?p_l_id=10718&folderId=288048&name=DLFE-27601.pdf)
- SACE Board of South Australia. (2011h). *Stage 2 2011 Clarifying forums evaluation: Final report*. Wayville: SACE Board of South Australia.
- SACE Board of South Australia. (2011i). *Stage 2 Clarifying forum booklet*. Wayville: SACE Board of South Australia.
- SACE Board of South Australia. (2011j). Teacher focus: Stage 2 Clarifying Forums. Retrieved 25 April, 2011, from <http://www.sace.sa.edu.au/>
- SACE Board of South Australia. (2012a). *Annual Report 31 December 2012*. Wayville: SACE Board of South Australia.
-



- SACE Board of South Australia. (2012b, July 2012). First Year Evaluation of the South Australian Certificate of Education: Final Report. Retrieved July 26, 2012, from [http://www.sace.sa.edu.au/\\_w/doc/SACE\\_First\\_Year\\_Evaluation\\_Report.pdf](http://www.sace.sa.edu.au/_w/doc/SACE_First_Year_Evaluation_Report.pdf)
- SACE Board of South Australia. (2012c). Physics 2011 Assessment Report. Retrieved 12 March, 2012, from <http://www.sace.sa.edu.au/documents/652891/0fd09867-0f8d-4378-af95-575975844603>
- SACE Board of South Australia. (2012d). Protocols and Procedures for Assessment Related Grievances, Appeals and Investigations. Retrieved 27 April, 2014, from <http://www.sace.sa.edu.au/documents/652891/92707cb4-5c8f-443c-a829-7dff833a539?v=1>
- SACE Board of South Australia. (2012e). SACE Assessment and Quality Assurance of Board-Accredited Subjects Policy. Retrieved 7 November, 2012, from <http://www.sace.sa.edu.au/documents/10336/0b963a6f-267d-4f1d-a50e-574d0ab4a534>
- SACE Board of South Australia. (2012f). SACE Assessment Responsibilities. Retrieved August 6, 2012, from [http://www.sace.sa.edu.au/home?p\\_p\\_id=20&p\\_p\\_lifecycle=1&p\\_p\\_state=exclusive&p\\_p\\_mode=view&\\_20\\_struts\\_action=%2Fdocument\\_library%2Fget\\_file&\\_20\\_folderId=343838&\\_20\\_name=DLFE-23301.pdf](http://www.sace.sa.edu.au/home?p_p_id=20&p_p_lifecycle=1&p_p_state=exclusive&p_p_mode=view&_20_struts_action=%2Fdocument_library%2Fget_file&_20_folderId=343838&_20_name=DLFE-23301.pdf)
- SACE Board of South Australia. (2012g). SACE First Year Evaluation: Evaluation Brief. Retrieved 21 December, 2012, from [http://www.sace.sa.edu.au/c/document\\_library/get\\_file?uuid=01ca3e12-c4fd-4cc8-a38f-9b9215a140d1&groupId=10336](http://www.sace.sa.edu.au/c/document_library/get_file?uuid=01ca3e12-c4fd-4cc8-a38f-9b9215a140d1&groupId=10336)
- SACE Board of South Australia. (2012h). Stage 2 Assessment and Quality Assurance. Retrieved 26 July, 2012, from [http://www.sace.sa.edu.au/c/document\\_library/get\\_file?p\\_l\\_id=343987&folderId=343838&name=DLFE-23302.pdf](http://www.sace.sa.edu.au/c/document_library/get_file?p_l_id=343987&folderId=343838&name=DLFE-23302.pdf)
- SACE Board of South Australia. (2012i). Stage 2 Physics Performance Standards. Retrieved 19 May, 2014, from <http://www.sace.sa.edu.au/web/physics/stage-2/support-materials/subject-advice-strategies>
- SACE Board of South Australia. (2013a). Annotated LAP exemplar. Retrieved 5 July, 2013, from <http://www.sace.sa.edu.au/web/physics/stage-2/support-materials/learning-and-assessment-plans>
- SACE Board of South Australia. (2013b). *Annual Report 31 December 2013*. Wayville: SACE Board of South Australia.
- SACE Board of South Australia. (2013c). Online Forums. Retrieved 2 September, 2013, from <http://www.sace.sa.edu.au/connect/online-forums>
- SACE Board of South Australia. (2013d). Procedures and Guidelines for Stage 2 Moderators—Final Moderation (Central). Wayville: SACE Board of South Australia.
- SACE Board of South Australia. (2013e). SACE Policy Framework. Retrieved 13 February, 2013, from <http://www.sace.sa.edu.au/documents/652891/705824/SACE+Policy+Framework.pdf>
- SACE Board of South Australia. (2014a). Assessment and Reporting for SACE Subjects. Retrieved 24 March, 2014, from

- <http://www.sace.sa.edu.au/documents/652891/d3530feb-007b-4dfe-aba5-93eb2c98e9ca?v=1>
- SACE Board of South Australia. (2014b). Assessment Task Exemplars. Retrieved 6 January, 2014, from <http://www.sace.sa.edu.au/web/physics/stage-2/support-materials/assessment-task-exemplars>
- SACE Board of South Australia. (2014c). Curriculum leaders groups. Retrieved 7 October, 2014, from <https://www.sace.sa.edu.au/about/consultation/curriculum-leadership-groups>
- SACE Board of South Australia. (2014d). Curriculum leaders groups. Retrieved 8 November, 2014, from <https://www.sace.sa.edu.au/about/consultation/curriculum-leadership-groups>
- SACE Board of South Australia. (2014e). Institute of Educational Assessors. Retrieved 18 October, 2014, from <https://www.sace.sa.edu.au/web/iea>
- SACE Board of South Australia. (2014f). Physics 2014 Subject Outline Stage 1 and Stage 2. Retrieved 6 January, 2014, from <http://www.sace.sa.edu.au/subjects/stage-2/sciences/physics>
- SACE Board of South Australia. (2014g). SACE Assessment Responsibilities. Retrieved 21 April, 2014, from <http://www.sace.sa.edu.au/documents/652891/b648bdb8-584f-46e5-a654-5f5b0e3a6496?v=1>
- SACE Board of South Australia. (2014h). SACE News Term 2 2014 May. Retrieved 14 May, 2014, from <http://www.sace.sa.edu.au/news/sace-news/2014/term-2-may#eoi>
- SACE Board of South Australia. (2014i, 2014). SACE Stage 2 Moderation Overview 2014. Retrieved 21 April, 2014, from <http://www.sace.sa.edu.au/documents/652891/6f854500-f764-4ad5-a892-51174ff9237d?v=1>
- SACE Board of South Australia. (2014j, 3 March 2014). Stage 2 Quality Assurance & Assessment Panels in 2014. Retrieved 2 October, 2014, from <https://www.sace.sa.edu.au/documents/652891/773541/Stage+2+Quality+Assurance+%26+Assessment+Panels.pdf/b11caa7d-7655-42e1-943a-2aa0a7a70e37>
- SACE Board of South Australia Act 1983*. (2010).
- SACE Principals Partnership Strategy Group. (2009). Meeting minutes 16 February 2009. Retrieved 28 April, 2014, from <https://www.sace.sa.edu.au/documents/10336/aea4c118-7087-4bd2-8dfe-011fe6a3cbf6>
- Sadler, D. R. (1987). Specifying and Promulgating Achievement Standards. *Oxford Review of Education*, 13(2), 191-209.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education*, 5(1), 77-84.
- Sadler, D. R. (2007). Perils in the meticulous specification of goals and assessment criteria. *Assessment in Education*, 14(3), 387-392.
- SATAC. (2012). Tertiary Entrance booklet. Retrieved 4 February, 2013, from

- <http://www.satac.edu.au/satac/TEB2013.pdf>
- Saxe, G., Gearhart, M., & Nasir, N. S. (2001). Enhancing students' understanding of Mathematics: A study of three contrasting approaches to professional support. *Journal of Mathematics Teacher Education*, 4, 55-79.
- Schon, D. A. (1987). *Educating the reflective practitioner*. San Francisco: Jossey-Bass.
- Schwartz-Shea, P., & Yanow, D. (2012). *Interpretive research design: concepts and processes*. New York, NY: Routledge.
- Shulman, L. S. (2002). Making Differences: A table of Learning. *Change: The Magazine of Higher Learning*, 34(6), 36-44.
- Smaill, E. (2013). Moderating New Zealand's National Standards: teacher learning and assessment outcomes. *Assessment in Education: Principles, Policy & Practice*, 20(3), 250-265.
- Spillane, J. P., Reiser, B., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, 72(3), 387-431.
- SSABSA. (2006). Assessment and Moderation Policy. Retrieved from <http://www.ssabsa.sa.edu.au/docs/policy/equity.pdf>
- Stanley, G., MacCann, R. G., Gardner, J., Reynolds, L., & Wild, I. (2009). Review of teacher assessment: evidence of what works best and issues for development. Retrieved 8 November, 2010, from [http://www.qcda.gov.uk/resources/assets/OUCEA\\_-\\_Review\\_of\\_teacher\\_assessment\\_March09.pdf](http://www.qcda.gov.uk/resources/assets/OUCEA_-_Review_of_teacher_assessment_March09.pdf)
- Starkey, L., Yates, A., Meyer, L. H., Hall, C., Taylor, M., Stevens, S. et al. (2009). Professional development design: Embedding educational reform in New Zealand. *Teaching and Teacher Education*, 25(1), 181-189.
- Suto, W. M. I., & Nádas, R. (2009). Why are some GCSE examination questions harder to mark accurately than others? Using Kelly's Repertory Grid technique to identify relevant question features. *Research Papers in Education*, 24(3), 335-377.
- Suurtamm, C., Koch, M., & Arden, A. (2010). Teachers' assessment practices in mathematics: classrooms in the context of reform. *Assessment in Education: Principles, Policy & Practice*, 17(4), 399-417.
- Swan, M. (2006). Designing and using research instruments to describe the beliefs and practices of mathematics teachers. *Research in Education*, 75, 58-70.
- Sykes, G. (1999). Teacher and Student Learning: Strengthening Their Connection. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: handbook of policy and practice* (1st ed., pp. 151-179). San Francisco, Calif.: Jossey Bass Publishers.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: HarperCollins College Publishers.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55.
- Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks: Sage Publications.

- Thompson, C. L., & Zeuli, J. S. (1999). The Frame and the Tapestry: Standards-Based Reform and Professional Development. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: handbook of policy and practice* (1st ed., pp. 341-375). San Francisco, Calif.: Jossey-Bass Publishers.
- Timperley, H. S. (2008). *Teacher professional learning and development*. Brussels: International Academy of Education.
- Timperley, H. S., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher Professional Learning and Development: Best Evidence Synthesis Iteration*. Wellington: New Zealand: Ministry of Education.
- University of Technology, Sydney. (2012). Students' approaches to learning. Retrieved 13 February, 2014, from <http://www.uts.edu.au/research-and-teaching/teaching-and-learning/learning-and-teaching/students-approaches-learning>
- van den Berg, R. (2002). Teachers' Meanings Regarding Educational Practice. *Review of Educational Research*, 72(4), 577-625.
- VCAA. (2008). VCE Physics Assessment Handbook 2009–2014. Retrieved 21 January, 2011, from [http://www.vcaa.vic.edu.au/vce/studies/physics/physics\\_hb.doc](http://www.vcaa.vic.edu.au/vce/studies/physics/physics_hb.doc)
- VCAA. (2013). VCE and VCAL Administrative Handbook, 2013. Retrieved February 22, 2013, from <http://www.vcaa.vic.edu.au/Documents/handbook/2013/AdHand2013.pdf>
- Webster-Wright, A. (2009). Reframing professional development through understanding authentic professional learning. *Review of Educational Research*, 79(2), 702-739.
- Wei, R. C., Darling-Hammond, L., & Adamson, F. (2010). *Professional development in the United States: Trends and challenges*. Dallas: National Staff Development Council.
- Wei, R. C., Darling-Hammond, L., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad*. Dallas, TX: National Staff Development Council.
- Wells, M. (2014). Elements of effective and sustainable professional learning. *Professional Development in Education*, 40(3), 488-504.
- Wenger, E. (1999). *Communities of practice: learning, meaning, and identity*. Cambridge: Cambridge University Press.
- Wiliam, D. (1992). Some technical issues in assessment: a user's guide. *British Journal of Curriculum and Assessment*, 2(3), 11-20.
- Wiliam, D. (2001). Reliability, validity, and all that jazz. *Education 3-13*, 29(3), 17-21.
- Willingham, W. W., Pollack, J. M., & Lewis, C. (2002). Grades and Test Scores: Accounting for Observed Differences. *Journal of Educational Measurement*, 39(1), 1-37.
- Wolf, A. (1995). *Competence-based assessment*. Buckingham: Open University Press.
- Wyatt-Smith, C. (1999). Reading for Assessment: How teachers ascribe meaning and value to student writing. *Assessment in Education: Principles, Policy & Practice*, 6(2), 195-223.
- Wyatt-Smith, C., & Castleton, G. (2005). Examining how teachers judge student writing: an Australian case study. *Journal of Curriculum Studies*, 37(2), 131–154.
- Wyatt-Smith, C., & Klenowski, V. (2013). Explicit, latent and meta-criteria: types of criteria

- 
- at play in professional judgement practice. *Assessment in Education: Principles, Policy & Practice*, 20(1), 35-52.
- Wyatt-Smith, C., Klenowski, V., & Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: a study of standards in moderation. *Assessment in Education: Principles, Policy & Practice*, 17(1), 59–75.
- Wylie, C., & Hodgen, E. (2010). NZCER 2010 Primary & Intermediate Schools National Survey: A snapshot of overall patterns and findings related to National Standards. Retrieved August 5, 2013, from <http://www.nzcer.org.nz/system/files/2010-primary-intermediate-national-survey.pdf>
- Yerrick, R., Parke, H., & Nugent, J. (1997). Struggling to promote deeply rooted change: The “filtering effect” of teachers’ beliefs on understanding transformational views of teaching science. *Science Education*, 81(2), 137-159.
- Young, M. F. D. (1998). *The curriculum of the future: from the "new sociology of education" to a critical theory of learning*. London: Falmer Press.
- Zemke, R., & Zemke, S. (1988). 30 Things We Know for Sure About Adult Learning. *Training*, 25(7), 57-61.
- Zemke, R., & Zemke, S. (1995). Adult learning: what do we know for sure? *Training*, 32(6), 31-40.
- Zepeda, S. J., Parylo, O., & Bengtson, E. (2013). Analyzing principal professional development practices through the lens of adult learning theory. *Professional Development in Education*, 40(2), 295-315.
- Zepke, N., Leach, L., Brandon, J., Chapman, J., Neutze, G., Rawlins, P. et al. (2005). *Standards-based assessment in the senior secondary school: A research synthesis. Final report to the New Zealand Qualifications Authority*. Palmerston North Massey University College of Education.